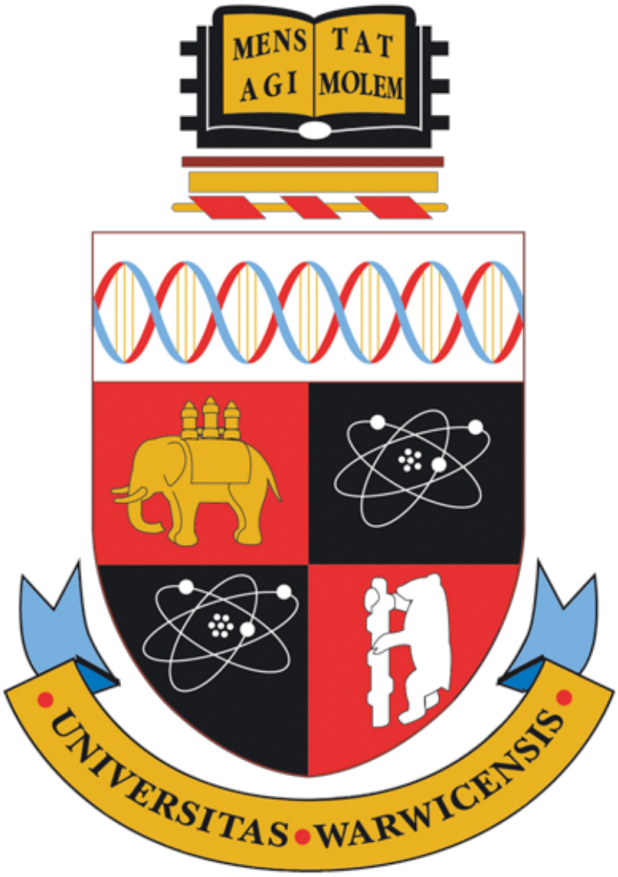
Functional genomics of methylated sulfur compound metabolism in *Hyphomicrobium* species



Richard Fetherston

This thesis is submitted for the degree of Doctor of Philosophy

School of Life Sciences, University of Warwick, Coventry, UK

September 2018

CONTENTS

LIST OF FIGURES

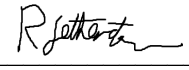
LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my supervisors Dr Hendrik Schäfer and Dr Yin Chen for their guidance and support through the PhD, my group members Dr Eileen Gröber, Jessica Palmer, Julie Scanlan and Dr Jason Stephenson for all their help in the lab and the rest of my colleagues, friends and family for their support. I'd also like to thank the NERC CENTA scheme for funding the PhD and my PhD advisors Prof Gary Bending and Dr Joseph Christie-Oleza for providing additional guidance. I'd particularly like to thank Dr Juan Ramon Hernández-Fernaud and Dr Cleidiane Zampronio at the Univerity of Warwick's Proteomics Facility for their technical assistance and Julie Scanlan for generating the *Hyphomicrobium sulfonivorans* mutant which went on to spur much of the research performed over the course of the project. Finally, I'd like to thank my examiners Prof Colin Murrel and Prof David Scanlan for their feedback on both the project and this thesis.

DECLARATION

This dissertation is the result of my own work and includes nothing, which is the outcome of work done in collaboration except where specifically indicated in the text. It has not been previously submitted, in part or whole, to any university or institution for any degree, diploma, or other qualification.

Signed: 

Date: 16/04/19

Richard Fetherston

ABSTRACT

Methylated sulfur compounds (MSCs) are a group of organosulfur compounds with various roles in agriculture, industry, medicine and the biogeochemical sulfur cycle, making their microbial metabolism an important area of study. The overarching aim of this project has been to investigate the molecular mechanisms allowing bacterial species to metabolise MSCs by studying members of the methylotrophic *Hyphomicrobium* genus - namely *Hyphomicrobium denitrificans* ATCC51888, *Hyphomicrobium methylovorum* Bras1 and *Hyphomicrobium sulfonivorans* S1 - using a combination of genetics, genomics, proteomics and transcriptomics. These model Alphaprotoebacteria can utilise MSCs such as dimethylsuilfide (DMS), dimethylsulfoxide (DMSO), dimethylsulfone ($DMSO_2$), methanesulfonic acid (MSA) and methanethiol (MT) as a sole source of carbon, sulfur and/or energy.

At the project's outset the mechanism *Hyphomicrobium* species use to degrade MSCs was thought largely understood, following the characterisation of a DMS oxidising DmoAB-type DMS monooxygenase from *Hyphomicrobium sulfonivorans* S1 and a MT oxidising MtoX-type MT oxidase from *Hyphomicrobium* sp. VS, leading to the suggestion that *Hyphomicrobium* species first reduce $DMSO_2$ and DMSO to DMS, then oxidise that DMS to hydrogen sulfide and formaldehyde via a MT intermediate.

However, this study now indicates that *Hyphomicrobium* species actually contain two distinct pathways of methylotrophic MSC metabolism. The first is a putative $DMSO_2$ oxidation pathway – proposed to contain the DmoAB and an SfnFG-type $DMSO_2$ monooxygenase – that generates formaldehyde and sulfite via the intermediate MSA leading to sulfite adenylylation for sulfur excretion. The second is a putative DMS oxidation pathway – proposed to contain MtoX and an as yet-unidentified DMS oxidising enzyme – that generates and formaldehyde and hydrogen sulfite via a MT intermediate leading to thiosulfate oxidation via the SOX system. Furthermore, this study identifies an additional non-methyltrophic $DMSO_2$ oxidation pathway in *H. sulfonivorans* that is specifically induced in response to sulfur starvation and capable of utilising $DMSO_2$, DMSO, DMS and MSA as a sole sulfur source, even in the absence of a functional DmoAB-type DMS monooxygenase.

ABBREVIATIONS

µg	microgram
µL	microliter
µM	micromolar
ABP	adenosin bisphosphate
ADP	adenosine diphosphate
APS	adenyl sulfate
ATP	adenosine triphosphate
bp	base pair
BLAST	basic local alignment search tool
BSA	bovine serum albumin
CBB	Calvin-Benson-Bassham
CCN	cloud condensation nuclei
CoA	coenzyme A
CoM	coenzyme M
Da	dalton
DMS	dimethylsulfide
DMSC	dimethyl sulfur compound
DMSO	dimethylsulfoxide
DMSO2	dimethylsulfone
DMSP	dimethylsulfoniopropionate
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DP	diphosphate
EC number	enzyme commission number
EDTA	ethylenediaminetetraacetic acid
FDR	false discovery rate
FID	flame ionisation detector
g	gram, gravitational force
GC	gas chromatography
GSH	glutathione

GRX	glutaredoxin
h	hour
H_4F	tetrahydrofolate
H_4MPT	tetrahydromethanopterin
hdr	heterodisulfide
kDa	kilodalton
KEGG	Kyoto Encyclopedia of Genes and Genomes
L	mitre
LC-MS	liquid chromatography–mass spectrometry
M	Molar
MeOH	methanol
mg	milligram
mL	millilitre
MS	mass spectrometry
MSA	methanesulfonate/ methanesulfonic acid
MSC	methylated sulfur compound
MSIA	methanesulfinate/ methanesulfinic acid
MT	methanethiol
NCBI	National Centre for Biotechnology Information
PAP	phosphoadenylyl sulfate
PCR	polymerase chain reaction
OSLH	O-succinyl-S-homoserine
RiBP	ribulose bisphosphate
RNA	ribonucleic acid
RNAseq	ribonucleic acid sequencing
rRNA	ribosomal RNA
RuMP	ribulose monophosphate
s	seconds
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
TRX	thioredoxin
VOSC	volatile organic sulfur compounds

Chapter 1:

Introduction

1. Introduction

1.1 Methylated sulfur compounds

Methylated sulfur compounds (MSCs) are a family of organosulfur compounds characterised by the presence of a methylated sulfur atom, such as methylated thiols, sulfides, sulfonium ions, sulfoxides or sulfones. Although MSC is sometimes used to denote all compounds containing a methylated sulfur atom, such as dimethylsulfoniopropionate, here MSC will only be used to refer to C1 methylated sulfur compounds that do not contain carbon-carbon bonds.

Some examples of common MSCs that are abundant in the marine and terrestrial environments are dimethylsulfide (DMS), dimethylsulfoxide (DMSO), dimethylsulfone ($DMSO_2$), methanethiol (MT), methane sulfonic acid (MSA) and methane sulfinic acid (MSIA), which are displayed in Figure 1.1-1. Although these compounds are similar in structure they display a diverse range of chemical properties; DMS and MT are volatile with a foul odour, DMSO and $DMSO_2$ are aprotic solvents, while MSA and MSIA are both protic solvents and strong acids.

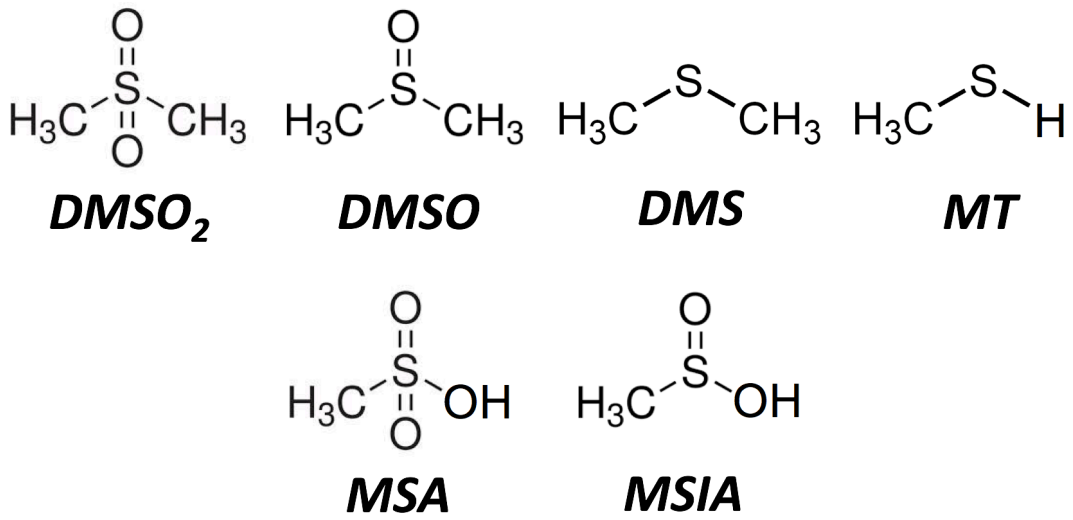


Figure 1.1-1: *Chemical structures of common methylated sulfur compounds*. Dimethylsulfone ($DMSO_2$), dimethylsulfoxide (DMSO), dimethylsulfide (DMS) and methanethiol (MT), methane sulfonic acid (MSA) and methane sulfinic acid (MSIA).

1.2 Biological, chemical and environmental relevance of MSCs

1.2.1 Biogeochemical cycling

Methylated sulfur compounds, specifically DMS, play a major role in the global sulfur cycle where they facilitate the transfer of sulfur from the oceans into the terrestrial environment (Curson *et al.*, 2011). This process centres around the production of DMS from the cleavage of dimethylsulfoniopropionate (DMSP), a highly abundant anti-oxidant, osmoprotectant and chemoattractant in the marine environment which plays an important role in the marine food web (Kiene *et al.*, 2000; Seymour *et al.*, 2010; Yoch, 2002).

This DMS-fated synthesis of DMSP in the oceans is often performed by sulfate fixing phototrophic algae – predominantly Dinophyceae and Prymnesiophyceae such as the model coccolithophore *Emiliania huxleyi* - for use as an osmolyte, representing as a much as 10% or more of the total organic carbon in marine phytoplankton (Keller *et al.*, 1989; Matrai and Keller, 1993; Matrai and Keller, 1994). However, as DMSP synthesis has also been discovered in certain species of heterotrophic marine bacteria, mostly Rhodobacterales but also some other species of Alphaproteobacteria, is also likely to represent a significant source of marine DMSP (Curson *et al.*, 2017). Such DMSP from either source is then released into the environment by the autolysis, viral lysis, senescence and grazing of these DMSP-containing microbes (Evans *et al.*, 2002). An estimated ~1 billion tonnes of DMSP is producted each year and the compound represents a substantial source of carbon and sulfur for other marine microorganisms (Curson *et al.*, 2011), especially as a metagenomic analysis performed by Howards *et al.* (2008) suggests that the majority of marine prokaryotes are capable of DMSP degradation.

The bacterial uptake and degradation of DMSP is a major source of environmental MSCs, mediated by a diverse range of enzymes using two distinct mechanisms; either the cleavage of DMSP by DMSP lyases to produce DMS (Curson *et al.*, 2008), or the sequential demethylation and demethiolation of DMSP to yield MT (Kiene, 1996). Although DMS, MT and the other DMSP metabolites are often used by these microbes as a carbon, sulfur and energy source, the sheer volume of bacterial MSC production in the marine environment has a substantial influence on the biosphere due in part to the chemical properties of DMS.

As a volatile organosulfur compound (VOSC), dissolved DMS is able to evaporate from the oceans into the atmosphere, leading to the release of ~10-30 million tonnes of sulfur each year and representing approximately 50% of total DMS emmisions into the atmosphere (Ayers and Gillet, 2000; Kettle and Andreae, 2000; Watts, 2000). However, it should be noted this marine DMS flux still only accounts for 8-10% of the DMS produced by the marine

environment each year, with the majority of DMS turned over by chemical and biological processes into sulfate, DMSO or incorporated into host biomass (Erikson *et al.*, 1990; Kiene and Linn, 2000).

The atmospheric and aqueous phase oxidation of DMS leads to the production of various aerosols and aqueous phase inorganic compounds including sulfur dioxide (SO_2), sulfuric acid (H_2SO_4) and sulfate (HSO_4 and SO_4^{2-}), and MSCs such as DMSO, $DMSO_2$ and MSA (Charlson *et al.*, 1987; Hoffman *et al.*, 2016). DMS and its oxidation products are then deposited into marine and terrestrial environments through precipitation, providing a biogeochemical pathway for the transfer of sulfur from the oceans onto the land.

Several of these oxidation products, namely H_2SO_4, SO_2 and MSA, can act as cloud condensation nuclei (CCN) to mediate cloud seeding, giving DMS at least some impact on oceanic weather patterns (Yoch, 2002; Korhonen *et al.*, 2008). In the past it was thought that DMS may participate in a climate-regulating negative feedback loop, commonly referred to as the CLAW hypothesis, whereby the microbial production of DMS leads to increased albedo from cloud condensation, a reduction in phytoplankton blooms and ultimately a decrease in DMS output (Charlson *et al.*, 1987).

However, as discussed by Quinn and Bates (2011), new developments in the field of atmospheric chemistry, climate modelling, and several decades of observation have yet to demonstrate such a substantial link between microbial DMSP production and climate, largely discrediting Charlson *et al.*'s original hypothesis (1987) though oceanic DMSP remains a key source of MSCs for both the marine and terrestrial environments, as well as an important intermediate of global sulfur cycling.

1.2.2 Other environmental and anthropocentric MSC sources

Although a large proportion of the world's MSCs originate from phototrophs in the marine environment, ocean derived DMS is only estimated to make up ~80% of the total DMS in the Earth's atmosphere (Watts, 2000). The remaining 20% is thought to be produced by a combination of plants, soils, salt marsh sediments and human activity (Watts, 2000; Schäfer 2009).

An example of this is the microbial production of DMS and MT from salt marshes, which are relatively large contributors of MSCs compared to other habitats (Steudler and Peterson, 1984; Kiene *et al.*, 1987; Kiene, 1988). Here, plant-derived DMSP is used as a

substrate for microbial growth in brackish salt marsh water, sediments and plant rhizospheres. This leads to the generation of DMS and MT via DMSP demethylation and lysis, in an analogous process to marine DMSP degradation discussed above (Yoch, 2002).

However, terrestrial MSC production is not only limited to salt marshes, as DMS and MT emission have also been detected from the plants and soil samples from agricultural land, grassland and woodland (Geng and Mu, 2004; Yi *et al.*, 2013; Liu *et al*., 2017) and there is some evidence to suggest that DMS production from the soil and vegetation in the Amazon Basin plays an important role in the region's microclimate (Jardine *et al.*, 2014). Furthermore, a new mechanism for bacterial DMS production has recently been described by Carrión *et al*. (2015) that appears to be ubiquitous in soil bacteria, in which methionine derived MT is methylated to DMS, and has been proposed as a major mechanism of terrestrial DMS production.

An overview of the major mechanisms of environmental MSC production, as well as the transfer of MSCs from the oceans into the terrestrial environment, is displayed in Figure 1.2-1.

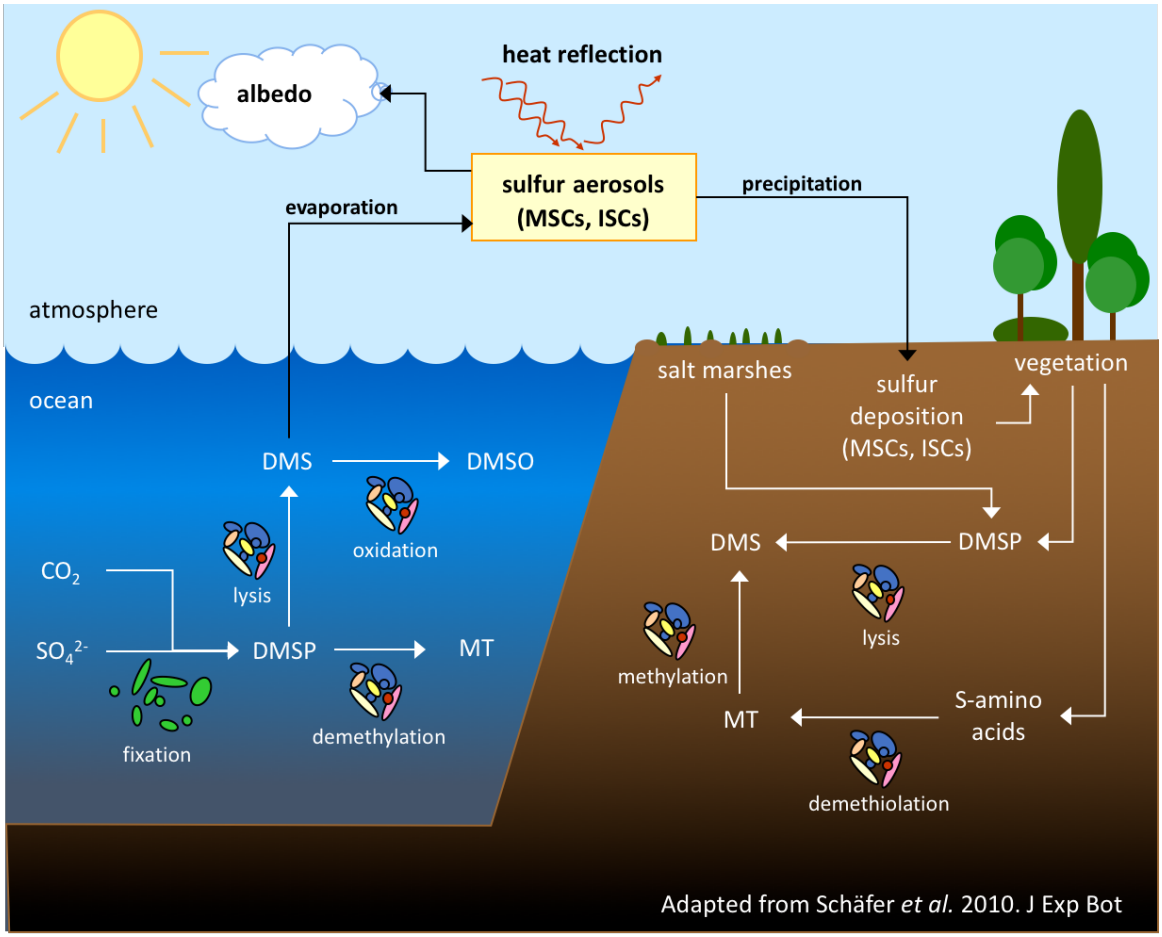


FIGURE 1.2-1: *Common sources of MSCs in the natural environmental.* (LEFT) Oceanic sulfate fixation to DMSP, leading to MSC formation by microbial DMSP cleavage and demethylation. (TOP) Atmospheric chemistry of DMS leading to the production of MSC and ISC (inorganic sulfur compound) sulfur aerosols, heat reflection and cloud seeding. (RIGHT) Deposition of MSCs and inorganic sulfur compounds (ISCs) into the terrestrial environment. Generation of DMSP by vegetation and salt marshes. Terrestrial production of DMS and MT from DMSP and sulfur containing amino acids (S-amino acids). Physical and chemical processes are displayed in BLACK, biological processes are displayed in WHITE.

1.2.3 Biofiltration, chemoattraction and malodour

The anthropocentric sources of MSCs include agriculture, domestic waste, industrial waste and in modest levels, human excretion and exhalation (Watts, 2000; Perraud, 2011; Giri *et al.*, 2015; Yong *et al.*, 2017). Due to the foul odour and low odour detection threshold of volatile MSCs, such as DMS and MT, they can be categorised as air pollutants when intersecting with human populations, often making them prospective targets for biofiltration (Ruokojärvi *et al.*, 2000; Fernández *et al.*, 2012; Giri *et al.*, 2015). Volatile MSCs also contribute to the aroma of fruit, vegetables and other food produce, as well as playing an important role as a chemoattract for animals and microbes in the marine environment (Nevitt, 2000; Yoch, 2002; Parker, 2015; Cannon and Ho, 2018).

A pungent example of MSC-mediated chemoattraction and foul odour is demonstrated by the plant *Helicodiceros muscivoras*, colloquially known as “dead horse arum lily” (Stensmyr *et al.*, 2002). This flowering plant synthesises the volatile MSCs dimethylsulfide, dimethyldisulfufide and dimethyltrisulfide, mimicking the odour of a rotting animal carcass to attract insects for pollination.

The study of malodourous MSCs also has some clinical relevance in oral health, in which DMS and MT are used as a clinical marker of both halitosis (bad breath) and general disease (Awano *et al.*, 2004; Tangerman and Winkel, 2012). Furthermore, the study of bacterial MSC metabolism has unexpectedly led to new developments in the human genetics of disease, in which a mutation of the human MT oxidase (SELENBP1), originally characterised in *Hyphomicrobium* species as MtoX (Eyice *et al.*, 2017), has been linked to both intraoral and extraoral halitosis (Pol *et al.* 2018).

1.3 Bacterial metabolism of methylated sulfur compounds

Numerous bacterial species have been isolated from the marine and terrestrial environments that are capable of utilising MSCs such as DMS, DMSO and MSA, as a carbon source, sulfur source and/or energy source for bacterial growth, representing a substantial sink for environmental MSCs (Kelly and Murrell, 1999; Schäfer *et al.*, 2009; Watts *et al.*, 2000). The various microbial processes of both methylotrophic and non-methylotorphic MSC metabolism are discussed below.

1.3.1 Methylotrophic MSC metabolism

A large number of the MSC utilising bacterial isolates identified to date are capable of utilising DMS and/or MT as a sole carbon and energy source (Schäfer *et al.*, 2009). This includes isolates from many of the DMS and MT producing environments discussed above in Section 1.2, such as freshwater, seawater, soil, salt marsh sediment, plant rhizospheres, activated sludge, coastal microbial mats and industrial biofilters (Sivelä and Sundman, 1975; Zwart *et al.*, 1996; Lomans *et al.*, 1999; Schäfer *et al.*, 2009, Eyice and Schäfer, 2016).

Examples have also been found of methylotrophic bacteria that are capable of using other MSCs as a sole carbon source, such as MSA degrading *Marinosulfonomonas methylotropha* and *Filomicrobium* species isolated from the marine environment (Thompson *et al.*, 1995; Baxter *et al.*, 2002; Henriques and De Marco, 2015), DMSO degrading *Hyphomicrobium* species from soil and activated sludge (De Bont *et al.*, 1981; Suylen and Kuenen, 1986; Eyice and Schäfer, 2016), *Hyphomicrobium* and *Variovorax* species from plant rhizospheres (Eyice and Schäfer, 2016), and $DMSO_2$ degrading *Arthrobacter* and *Hyphomicrobium* species isolated from soil (Borodina *et al.*, 2000; Kino *et al.*, 2004). It may be interesting to note that the methylotrophic metabolism of DMSO and $DMSO_2$ appears to be more prevalent in isolates from the terrestrial environment than the marine environment (Schäfer *et al.*, 2009), but again this may be due to the sampling bias of previous isolation experiments.

1.3.2 Non-methylotrophic MSC metabolism

Although fewer enrichment experiments have been performed for bacterial utilisers of MSCs as a sole sulfur or energy source than for methylotrophic MSC utilisation, various

bacterial species have been identified that are capable of utilising MSCs as a sole sulfur source but not as a sole carbon source. Examples include the MSA utilisers *Bacillus subtilis* and *Escherichia coli* (Ploeg *et al.*, 1998; Eichhorn and Leisinger, 2001), the DMS utilising marine isolate *Acinetobacter guillouiae* (Horinouchi *et al.*, 1997) and several $DMSO_2$, DMSO, DMS and MSA utilising *Pseudomonas* species (Ploeg *et al.*, 1998; Kahnert *et al.*, 2000; Endoh *et al.*, 2003).

One of the most interesting isolate phenotypes of MSC utilisation is that of *Sagittula stellata* obtained from the marine environment, capable of utilising DMS as an energy source by the oxidation of DMS to DMSO, but incapable of utilising the compound as either a carbon or sulfur source (Boden *et al.*, 2011b).

1.3.3 Pathways and products of MSC metabolism

Although there is no ubiquitous mechanism of bacterial MSC metabolism, both methylotrophic and non-methylotrophic MSC metabolism commonly involves their oxidation or demethylation to yield formaldehyde and/or a methylated cofactor, and either hydrogen sulfide or sulfite (Schäfer *et al.*, 2009). The assimilatory and dissimilatory metabolism of these carbon and sulfur products are outlined in Section 1.4 and 1.5 respectively, while an overview of the various enzymes of MSC metabolism is described below.

So far, three bacterial enzymes have been implicated in the methylotrophic metabolism of MSCs: a DmoAB-type DMS monooxygenase characterised in *H. sulfonivorans* S1 (Boden *et al.*, 2011), a MT oxidase characterised in *Hyphomicrobium* species VS which oxidises MT (Eyice *et al.*, 2017) and an MsmABCD-type MSA monooxygenase from *Marinosulfonomonas methylotropha* (Thompson *et al.*, 1995; Baxter *et al.*, 2002).

Two of the enzymes responsible for mediating non-methylotrophic MSC metabolism in these species show functional overlap with the MsmABCD-type MSA monooxygenase of methylotrophic MSC metabolism: the MsuDE-type MSA monooxygenase has been characterised in *Pseudomonas putida* (Kertesz *et al.*, 1999) and an SsuDE-type alkanesulfonate monooxygenase found in *B. subtilis* and *E. coli* (Ploeg *et al.*, 1998; Eichhorn and Leisinger, 2001). A related enzyme for non-methylotrophic MSC utilisation has also been identified in *Pseudomonas fluorescens*, SfnFG-type $DMSO_2$ monooxygenase, which oxidises $DMSO_2$ with an electron donor to produce formaldehyde and MSIA (Wicht, 2016).

Two examples have also been found for enzymes catalysing the oxidation of DMS to the related dimethyl sulfur compound DMSO, the respective DsoABCDEF-type DMS oxidation system of *Acinetobacter guillouiae* (Horinouchi *et al.*, 1997; Horinouchi *et al.*, 1999) and DdhABC-type DMS dehydrogenase of *Rhodovulum sulfidophilum* (McDevitt *et al*, 2002). Various enzyme assays performed in *Hyphomicrobium* and *Pseudomonas* species have also found enzyme activity relating to the reduction of $DMSO_2$ to DMS, and the oxidation of DMS to $DMSO_2$ (Borodina *et al.*, 2000; Boden *et al.*, 2011; Endoh *et al.*, 2003). It is thought that these may represent a form of dimethyl sulfur compound dehydrogenase and/or reductase, but as they have yet to be fully characterised and identified in their respective organisms they still remain somewhat hypothetical.

A simplified scheme of bacterial MSC catabolism is displayed in Figure 1.3-1.

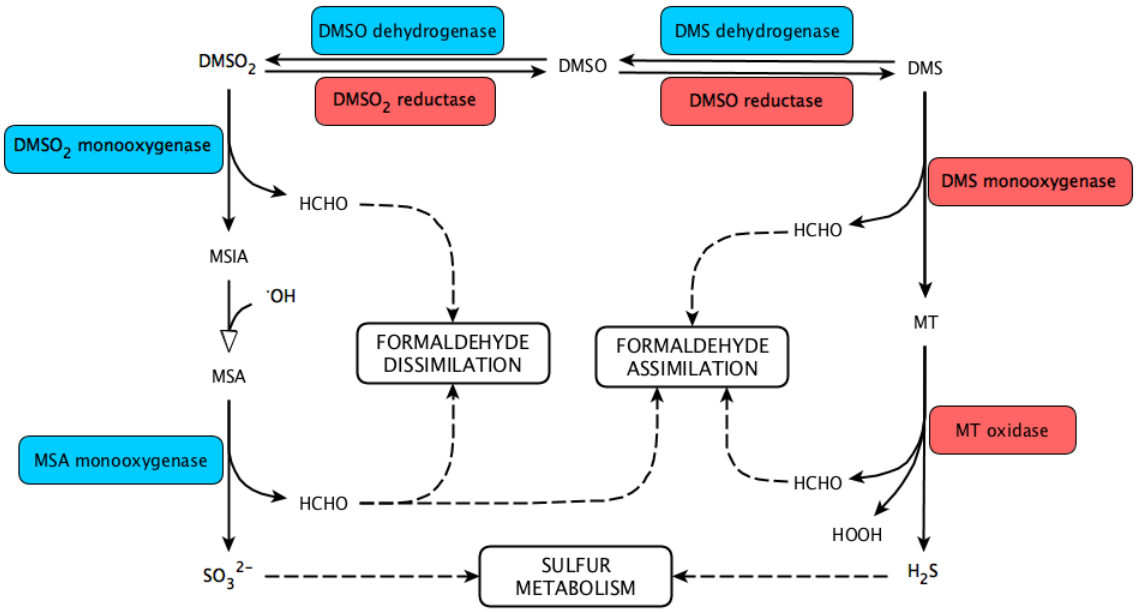


Figure 1.3-1: *Bacterial mechanisms of catabolic MSC metabolism.* Oxidation of MSCs to sulfite via $DMSO_2$ is displayed in BLUE. The catabolism of MSCs to hydrogen sulfide via DMS is displayed in RED. The known fates of formaldehyde (HCHO) from these processes are displayed in WHITE. Note that this figure only shows the potential mechanisms of bacterial MSC metabolism as they are currently understood and does not represent of MSC metabolism in any one organism.

1.4 Methylotrophic carbon metabolism

Methylotrophy is the capacity to utilise C1 compounds as a carbon and energy source; C1 refers to carbon containing compounds without carbon-carbon bonds such as methane, methanol, methylated amides or methylated sulfides (Chistoserderva *et al.*, 2009; Chistoserderva, 2011). Facultative methylotrophs are able to use both C1 compounds and organic compounds containing carbon-carbon bonds as a sole carbon and sulfur source, whereas obligate methylotrophs are only able to use C1 compounds.

A common mechanism of methylotrophy in prokaryotes is the oxidation of C1 compounds to yield the intermediates formaldehyde, formate and/or carbon dioxide for subsequent utilisation as a carbon and/or energy source (Chistoserderva *et al.*, 2009; Chistoserderva, 2011). Formaldehyde is a ubiquitous, toxic product of demethylation reactions in living organisms that has non-specific reactivity with nucleic acids and proteins (Chen *et al.*, 2016), making formaldehyde detoxification an essential process for methylotrophic organisms (Chongcharoen *et al.*, 2005). Depending on the C1 compound involved this may either be a simple single-step process mediated by a single enzyme, such as MeOH oxidation (Schmid *et al.*, 2010) or a series of reactions involving multiple enzymes and generating a variety of different metabolites, such as the methylotrophic metabolism of trimethylamine (Colby and Zatman, 1973).

Common types of assimilatory C1 enzymes are the electron-acceptor dependent dehydrogenases, oxygen-dependent oxidases, oxygen and electron donor-dependent monooxygenases, and methyl acceptor-dependent methyltransferases. For example, methanol dehydrogenase uses NAD^+ or $NADP^+$ as an electron acceptor to oxidise methanol to formaldehyde (Schmid *et al.*, 2010); MT oxidase oxidises MT with molecular oxygen and water to produce formaldehyde, hydrogen sulfide and hydrogen peroxide (Eyice *et al.*, 2017); particulate methane monooxygenase uses molecular oxygen and the artificial electron donor quinol to produce methanol, quinone and water (Myronova *et al.*, 2005). Note that unless the methyl acceptor is itself an assimilation intermediate, such as the chloromethane:tetrahydrofolate methyltransferase seen in *Methylobacterium chloromethanicum*, then the methyl group will need to be removed from the acceptor by a secondary mechanism for use as a carbon source (Studer *et al.*, 2001).

Methylotrophic bacteria have historically been sorted into one of three categories, broadly styled after the molecular mechanisms discovered in three model organisms – the Type I methylotroph *Methylomonas methanica,* Type II methylotroph *Methylosinus trichosporium*

and the type X methylotroph *Methylococcus capsulatus* (Chistoserdova *et al.*, 2009). *Type* I methylotrophs assimilate formaldehyde via the ribulose monophosphate (RuMP) pathway, while Type II methylotrophs assimilate formaldehyde via the serine cycle, and Type X methylotrophs primarily assimilate formaldehyde via the ribulose monophosphate (RuMP) cycle, but also contain enzymes from the serine cycle (Chistoserderva *et al.*, 2009; Chistoserderva, 2011).

Each cycle/pathway uses a different C1 or C1 derivative as its substrate: the ribulose monophosphate (RuMP) cycle to generate pyruvate for carbon assimilation, using amino acids and carboxylic acids as intermediates; the Calvin-Benson-Bassham (CBB) cycle uses the formaldehyde derivative carbon dioxide to generate glyceraldehyde-3-phosphate; the serine cycle uses the formaldehyde/formate derivative 5,10-methylene tetrahydrofolate to produce 3-phospho-D-glycerate, using amino acid intermediates (Vorholt, 2002; Chistoserdova, 2011).

Although the RuMP cycle utilises formaldehyde as a carbon source directly, further metabolism of formaldehyde is required to generate substrates of the CBB cycle and serine cycle, carbon dioxide and 5,10-methylene tetrahydrofolate respectively. In CBB cycle methylotrophs and some serine cycle methylotrophs, this process will involve the conversion of formaldehyde to formate (Chistoserderva *et al.*, 2009; Chistoserdova, 2011), typically mediated by one of three distinct mechanisms: The glutathione-dependent formaldehyde activation pathway uses glutathione (GSH) as a cofactor for formate generation, the tetrahydromethanopterin-dependent formaldehyde activation pathway (H_4MPT pathway) uses tetrahydromethanopterin (H_4MPT) and methanofuran (MFR) as cofactors to generate formate, while a formaldehyde dehydrogenase oxidises formaldehyde to formate using NAD^+ as an electron acceptor (Vorholt, 2002).

For assimilatory formate metabolism via the CBB cycle, formate is further oxidised by an NAD^+-dependent formate dehydrogenase to produce carbon dioxide, though it should be noted that this reaction also serves as a mechanism for the dissimilatory metabolism of formate to carbon dioxide.

The production of 5,10-methylene tetrahydrofolate for assimilation via the serine cycle may have two distinct mechanisms; either the direct condensation of tetrahydrofolate (H_4F) with formaldehyde, or the formate-dependent methylation of H_4F via the H_4F pathway (Vorholt, 2002; Crowther *et al.*, 2008). However, as the H_4F pathway is bidirectional, the direction of the pathway between assimilatory production or dissimilatory demethylation of 5,10-methylene H_4F will depend on the specific mechanism used in a given organism.

A simplified scheme of formaldehyde metabolism is displayed in Figure 1.4.1 and will be described in more detail in Chapter 4.

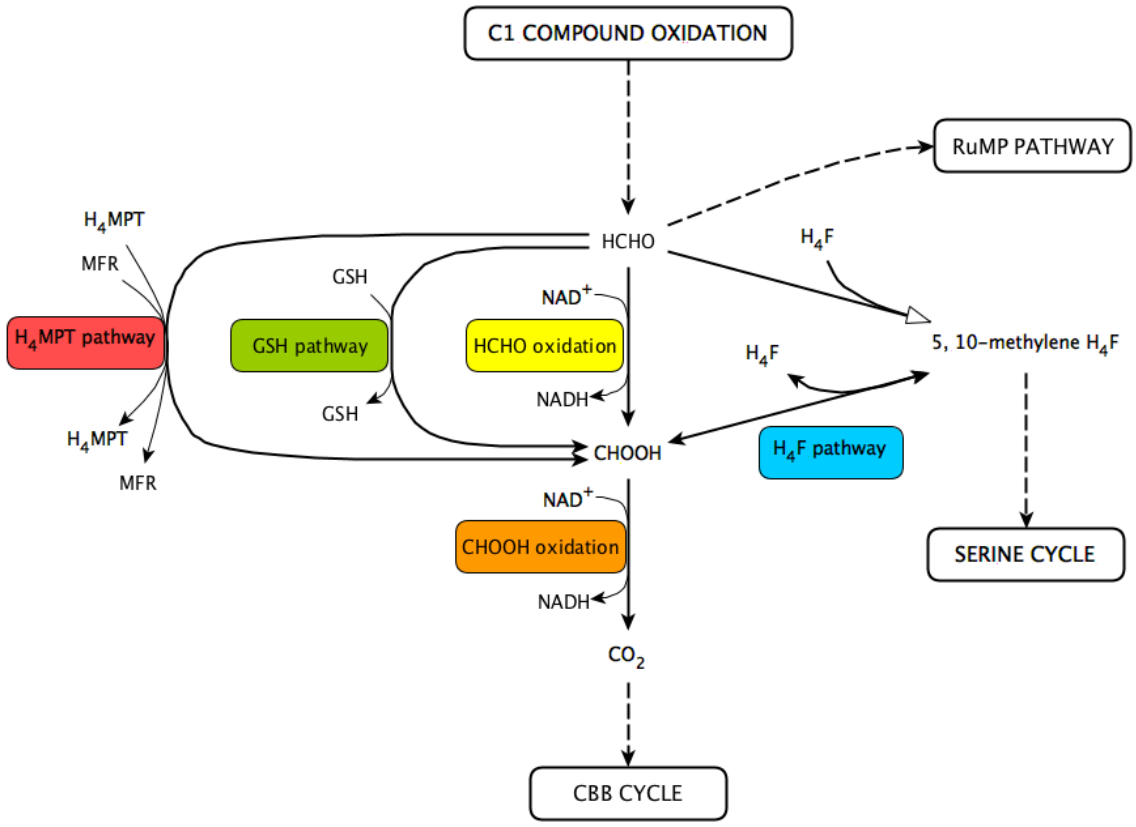


Figure 1.4-1: *Simplified scheme of bacterial assimilatory and dissimilatory formaldehyde metabolism.* Metabolic processes are labelled as follows: H_4MPT-dependent formaldehyde activation pathway (RED), GSH-dependent formaldehyde activation pathway (GREEN), formaldehyde oxidation (YELLOW), formate oxidation (ORANGE), H_4F pathway (BLUE), spontaneous condensation of formaldehyde with H_4F (WHITE arrowhead).

1.5 Inorganic sulfur metabolism

Sulfur is an essential element of life, predominantly taking the form of inorganic sulfur compounds such as sulfate, sulfite and thiosulfate, organosulfur compounds such as alkyl sulfides, sulfonates, sulfate esters or sulfur containing amino acids, cofactors and fatty acids (Sekowska *et al.*, 2000; Kertetsz, 2000). The metabolism and assimilation of inorganic sulfur compounds is therefore an important process for bacterial growth, consisting of a complex set of interlinked and overlapping metabolic pathways that broadly mediate three mechanisms: assimilatory sulfate reduction, excretory sulfide oxidation and amino acid synthesis. Although various other pathways of sulfur oxidation and reduction are also found in bacterial species, such as dissimilatory sulfate reduction by sulfate reducing bacteria (Santos *et al.*, 2015), the aforementioned mechanisms of sulfur metabolism described below represent the major paths of sulfur metabolism that are relevant to this study.

1.5.1 Assimilatory sulfate reduction

The bacterial utilisation of inorganic sulfur compounds as a sulfur source primarily occurs via the production of the sulfur containing amino acid cysteine, followed by its conversion to methionine, coenzyme A, glutathione, molybdopterin or various other organosulfur compounds (Kredich, 2008).

A common bacterial assimilation mechanism for inorganic sulfur compounds is to import extracellular sulfate and convert the sulfate to sulfite via a sulfate adenylylation pathway. Here, ATP is used to adenylylate sulfate to adenyl sulfate (adenosine 5'-phosphosulfate, APS) which is then phosphorylated with a second molecule of ATP to yield 3'-phosphoadenylyl sulfate (PAPS), and reduced to produce sulfite (Sekowska *et al.*, 2000; Kertetsz, 2006). The resultant sulfite is further reduced to sulfide by an NADH-dependent sulfite reductase, which can then be used as a substrate for amino acid synthesis.

The assimilation of sulfide as a sulfur source is typically mediated by a cysteine synthase, which reacts hydrogen sulfide with O-acetyl serine to produce L-cysteine and acetone (Kredich, 2008). Sulfur assimilation via homocysteine synthesis has also been observed in some bacterial species, in which sulfide is transferred onto O-succinyl-L-homoserine to produce succinate and L-homocysteine (Vermeij and Kertesz, 1999; Yoshida *et al.*, 2003), though cysteine synthesis is likely to be the dominant mechanism in most bacterial species (Kredich, 2008).

Various bacterial species, the most studied being *E. coli* and *Mycobacterium tuberculosis*, are also capable of utilising thiosulfate instead of sulfate as a substrate for assimilatory cysteine synthesis, often using the same enzyme and substrate to mediate the reaction (Sekowska *et al.*, 2000; Nakatani *et al.*, 2012; Steiner *et al.*, 2014); thiosulfate is reacted with O_3-acetyl-L-serine by a cysteine synthase to produce S-sulfocysteine, which can be reduced by a redoxin, such as thioredoxin or glutaredoxin, to generate L-cysteine and one molecule of sulfite.

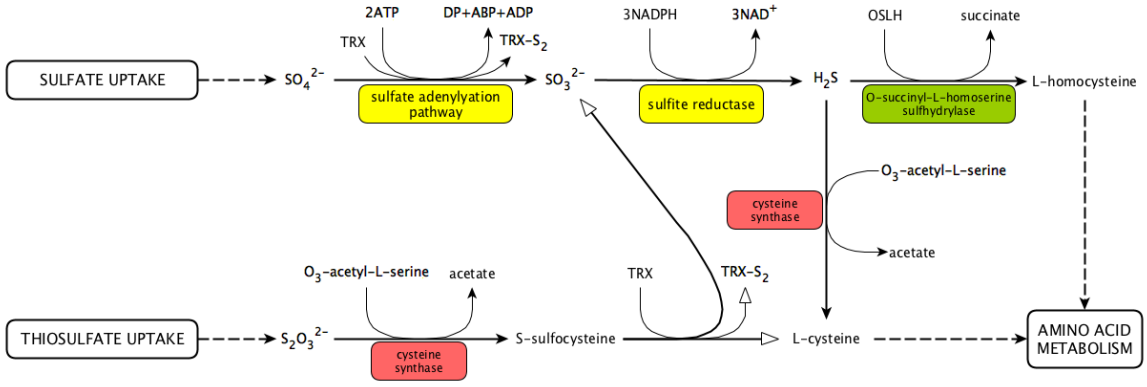


Figure 1.5-1: *Simplified scheme of assimilatory bacterial sulfur reduction and cysteine synthesis*. Reaction pathways for the reduction of sulfate (SO_4^{2-}) to hydrogen sulfide (H_2S) via sulfite (SO_3^{2-}) are labelled in YELLOW. Cysteine synthesis from hydrogen sulfide and thiosulfate mediated by cysteine synthase is labelled RED. L-homocysteine synthesis mediated by OSLH sulfhydrylase is labelled GREEN. Spontaneous reaction of S-sulfocysteine with TRX is labelled with WHITE arrowheads. Note the uncommon abbreviations TRX (thioredoxin), DP (diphosphate), ABP (adenosine bisphosphate) and OSLH (O-succinyl-L-homoserine).

1.5.2 Excretory sulfide oxidation

Although inorganic sulfur compounds are essential bacterial nutrients, found in the sulfur-containing amino acids (cysteine, homocysteine, methionine and taurine) and various other organosulfur compounds, both sulfite and hydrogen sulfide can be toxic to bacteria at high levels (Gunnison, 1981; Wang, 2012), necessitating a mechanism for their oxidation and excretion from the cell. Although certain species of bacteria are capable of excreting hydrogen sulfide and sulfite directly, it may be considered wasteful given the oxidation potential of these compounds for use as an energy source (Xia *et al.*, 2017). Several common oxidation mechanisms are described below and displayed in Figure 1.5-2.

Hydrogen sulfide can be oxidised to sulfite by the combined activity of three enzymes: a sulfur:quinone oxidoreductase, rhodanese and persulfide dioxygenase (Xia *et al.*, 2017). Sulfite can be further oxidised to sulfate directly by a sulfite dehydrogenase (Kappler and Dahl, 2001), or may occur via the oxidative adenylylation of sulfite following the reverse sulfate adenylylation pathway; via the oxidation of sulfite to PAPS, dephosphorylation to adenyl sulfate and deadenylylation to produce sulfate.

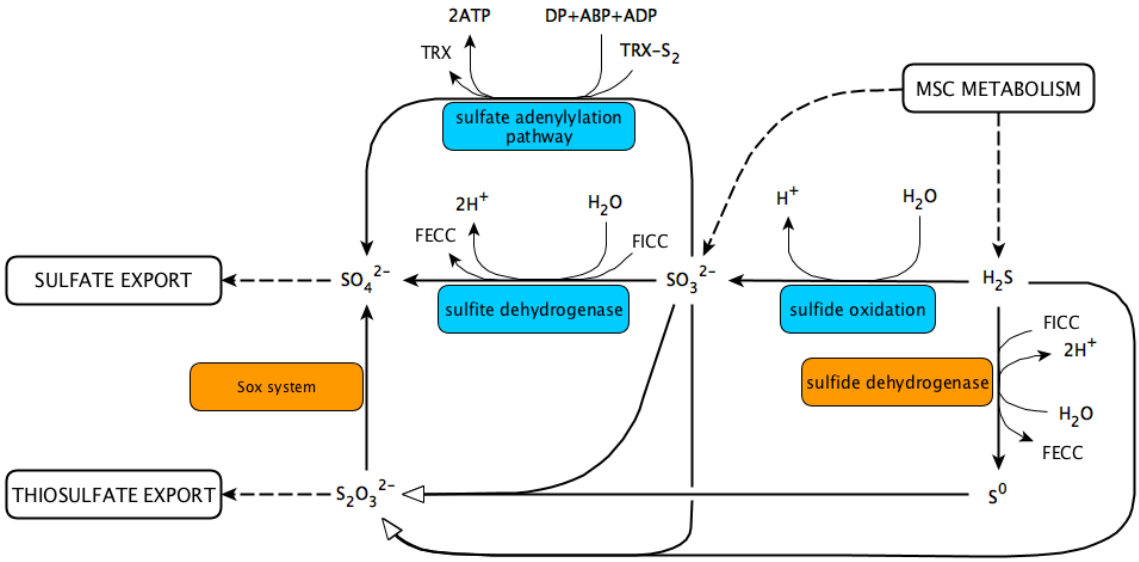


Figure 1.5-2: *Simplified scheme of excretory sulfide oxidation*. Reaction pathways for the oxidation of hydrogen sulfide (H_2S) to sulfate (SO_4^{2-}) via sulfite (SO_3^{2-}) are labelled in BLUE. Reaction pathways specific for the oxidation of hydrogen sulfide (H_2S) via thiosulfate ($S_2O_3^{2-}$) are labelled in ORANGE. Spontaneous reaction of sulfate with either hydrogen sulfide or sulfur (S^0) to form thiosulfate is labelled with WHITE arrowheads. Note the uncommon abbreviations TRX (thioredoxin), DP (diphosphate), ABP (adenosine bisphosphate), FICC (ferri cyctochrome *c*) and FECC (ferro cytochrome *c*).

Alternatively, sulfite can spontaneously react with sulfide, persulfide and/or sulfur (via a sulfide dehydrogenase) to generate thiosulfate, which can either be excreted from the cell or further oxidised by a thiosulfate oxidation system. A widespread example of a bacterial thiosulfate oxidation system is the near ubiquitous Sox system, which plays a major role in lithotrophic sulfur oxidation among the Alphaproteobacteria and has homologues in most other bacterial species (Ghosh and Dam, 2009).

The ‘complete’ Sox system SoxAXBCDYZ or SoxABCDXYZ consists of four enzymes with seven distinct subunits: SoxAX, SoxB, SoxCD and SoxYZ (Friedrich *et al.*, 2005). Together, these enzymes oxidise one molecule of thiosulfate to produce two molecules

of sulfate, plus four electrons and ten protons for the electron transport chain (Grabarczyk *et al.*, 2015). A reaction scheme for the Sox system from the model Alphaproteobacteria *Paracoccus pantrophus* is shown in Figure 1.5-3. Although doubts have been raised over the voracity of SoxYZ biochemistry in this particular model – SoxY may carry sulfate via R-S-S-$S_2O_3^-$ rather than R-S-$S_2O_3^{2-}$ (Grabarczyk and Berks, 2017) – it is an accurate enough approximation for the purposes of this study.

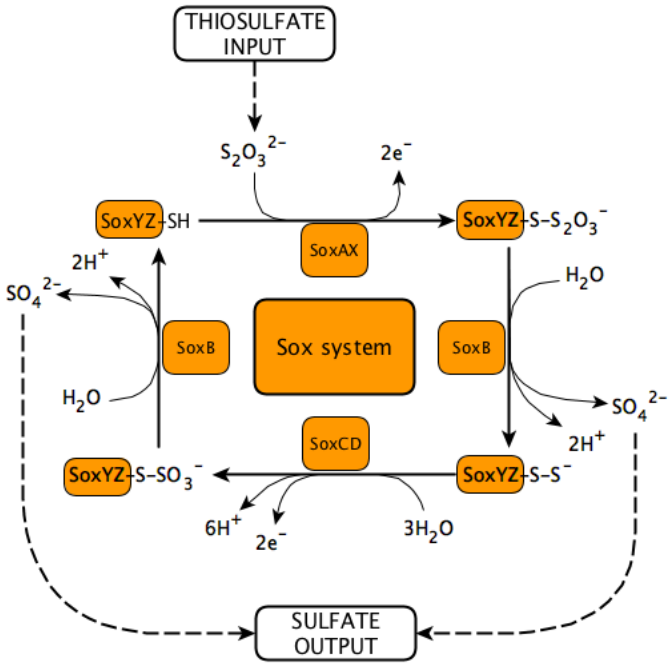


Figure 1.5-3: *Simplified scheme of the thiosulfate oxidising Sox system. SoxAXBCDYZ.* SoxAX conjugates thiosulfate onto a cysteine residue of SoxY (R-SH) to form two electrons and a sulfonate (R-S-$S_2O_3^-$), which is hydrolysed by thiosulfohydrolase SoxB to form sulfate (SO_4^{2-}) and a sulfane group (R-S-S^-). This is oxidised to a new sulfonate (R-SO_3^-) by the sulfane dehydrogenase SoxCD also producing two electrons, then hydrolysed by SoxB once more to produce a second molecule of sulfate and regenerate SoxYZ, beginning the cycle anew. Adapted from Grabarczyk *et al.* (2015).

1.6 *Hyphomicrobium*, a model genus of methylotrophic bacteria

Hyphomicrobium is a genus of methylotrophic Alphaproteobacteria from the family *Hyphomicrobiaceae*, chosen as model organisms for this project due to the capacity of several *Hyphomicrobium* species to utilise MSCs as a sole carbon, sulfur and energy source (De Bont *et al.*, 1981; Suylen and Kuenen, 1986). Morphologically they are ovoid, budding, non-spore forming, Gram-negative, prosthecate bacteria (Moore, 1981); prosthecates are filamentous appendages sprouting from the cell body that are believed to be cell membrane extensions to increase bacterial surface area for nutrient uptake (McAdams, 2006).

Although the *Hyphomicrobium* were described by Rullman as early as 1897 with the observation of *Hyphomicrobium vulgar*e in bacterial enrichment cultures, it was only in the 1960s that an effective method for their enrichment and isolation was developed by Hirsch and Conti, then later refined by Attwood and Harder in the 1970s (Hirsch and Conti, 1964; Atwood and Harder, 1972). Indeed, by the time of a review by Moore in 1981 the genus had already been observed in soil, freshwater, seawater, brackish water, sewage, acidic mine water and laboratory water baths (Moore, 1981).

Methylotrophy in the genus has been studied extensively over the past fifty years, with much of the work performed in *Hyphomicrobium* X, identifying the *Hyphomicrobium* as a genus of serine cycle methylotrophs which are capable of utilising a wide variety of different C1 compounds (Harder *et al.*, 1973; Harder and Attwood, 1975; Attwood and Harder, 1978). This includes formaldehyde, formate and methanol (Moore, 1981), methylated amines (Meiberg and Harder, 1977; Brooke *et al.*, 1987), chloromethanes (Doronina *et al.*, 1996; McAnulla *et al.*, 2001; McDonal *et al.*, 2001) and, importantly for this project, MSCs.

Several *Hyphomicrobium* species are capable of utilising MSCs as methylotrophic carbon sources. Both the soil isolate *Hyphomicrobium* S and waste water isolate *Hyphomicrobium* EG were found to be capable of utilising DMS and DMSO as a sole carbon, sulfur and energy source (De Bont *et al.*, 1981; Suylen and Kuenen, 1986; Suylen *et al.*, 1986). Later enrichment experiments have gone on to isolate several more MSC utilising *Hyphomicrobium* species, including *Hyphomicrobium sulfonivorans* (Borodina *et al.*, 2000), as well as several other *Hyphomicrobium* species that will be covered in later chapters.

1.7 Project aims and objectives

The overarching aim of the project has been to use functional genomics to improve our understanding of MSC metabolism in the model organism *Hyphomicrobium sulfonivorans* S1 and apply this information to the study of other *Hyphomicrobium* strains to gain a better understanding of the bacterial processes of assimilatory MSC metabolism. To this end, the project had four broad objectives:

1. **Characterise the capacity of *H. sulfonivorans* S1 to utilise MSCs**
2. **Identify the enzymes responsible for the catabolism of $DMSO_2$ in *H. sulfonivorans***
3. **Characterise the capacity of other *Hyphomicrobium* species to utilise MSCs**
4. **Compare and contrast MSC metabolism in different *Hyphomicrobium* species**

This work has been divided into three experimental chapters, beginning with the genotyping and phenotyping of *H. sulfonivorans* in Chapter 3, the functional genomics of *H. sulfonivorans'* MSC metabolism in in Chapter 4, and an examination of MSC utilisation in other *Hyphomicrobium* species using a combination of genotyping, phenotyping and comparative proteomics in Chapter 5.

Chapter 2:

Materials and Methods

2.1 Bacterial species and strains

Table 2.1.1: *List of bacterial strains, sources and first publication.*

Strain	Source	Publication
Hyphomicrobium denitrificans ATCC 51888	ATCC 51888	Harder *et al.,* 1973
Hyphomicrobium facile Bras3	Schäfer lab group	Eyice and Schäfer, 2016
Hyphomicrobium methylovorum Bras1	Schäfer lab group	Eyice and Schäfer, 2016
Hyphomicrobium sulfonivorans S1	DSMZ 13863	Bordina *et al.,* 2000
Hyphomicrobium sulfonivorans S1 Δ*dmoA*	Scanlan & Schäfer	unpublished
Hyphomicrobium species VS	Op den Camp, Nijmegen	Pol *et al.,* 1994

2.2 Growth media, substrates and supplements

2.2.1 Complete Basal Salts

Basal Salts (CBS) medium is a minimal growth media that was used for the cultivation of *Hyphomicrobium* species. A 1L solution of CBS media is produced in two parts, A and B. An 800 mL solution of CBS part A, contains 0.2g ammonium chloride, 0.1g $MgSO_4.7H_2O$ and 10 mL CBS NS trace metal solution.

A 200 mL solution of CBS part B, contains 1.5g KH_2PO_4 and 6.2g Na_2HPO_4. Both solutions were autoclaved at 121°C for 15 min, stored in the dark and combined under sterile conditions at RT to produce 1 L CBS NS media.

To produce a 1 L stock of CBS trace metal solution, 50 g of Na_2EDTA (ethylenediaminetetraacetic acid) and 10 g of NaOH are dissolved in 200 mL $d.H_2O$ (deionised water), followed by the sequential addition of 50 mL $d.H_2O$ solutions of each the following compounds: 11.5 g $ZnSO_4$, 7.34 g $CaCl_2.2H_2O$, 2.5 g $MnCl_2.4H_2O$, 0.5 g $CoCl_2.6H_2O$, 0.5 g ammonium molybdate, 0.2 g $CuSO_4.5H_2O$, 0.05 g boric acid, 0.01 g ammonium metavandate, 5 g $FeSO_4.7H_2O$. The stock is then adjusted to pH 6.0 by the addition of 1 M NaOH and brought to a final volume of 1 L with $d.H_2O$.

2.2.2 Complete Basal Salts – sulfur free (CBS NS)

CBS NS media is a minimal, sulfur free growth media that was used for the cultivation of *Hyphomicrobium* species, developed by substituting metal sulfates from CBS media for chlorides. Like CBS media, a 1 L solution of CBS NS media is produced in two parts, A and B. An 800 mL solution of CBS NS part A, contains 0.20 g ammonium chloride, 0.085 g $MgCl_2.H_2O$ and 10 mL CBS NS trace metal solution. A 200 mL solution of CBS NS part B,

contained 1.5 g KH_2PO_4 and 6.2 g Na_2HPO_4. Both solutions were autoclaved at 121°C for 15 min, stored in the dark and combined under sterile conditions at RT to produce 1 L CBS NS media.

To produce a 1 L stock of CBS NS trace metal solution for sulfur free media, 50 g of Na_2EDTA and 10 g of NaOH are dissolved in 200 mL d.H_2O, followed by the sequential addition of 50 mL d.H_2O solutions of each the following compounds: 9.71 g $ZnCl_2$, 7.34 g $CaCl_2.2H_2O$, 2.5 g $MnCl_2.4H_2O$, 0.5 g $CoCl_2.6H_2O$, 0.5 g ammonium molybdate, 0.15 g $CuCl_2.2H_2O$, 0.05 g boric acid, 0.01 g ammonium metavanadate, 3.56 g $FeCl_2.4H_2O$. The stock is then adjusted to pH 6.0 by the addition of 1 M NaOH and brought to a final volume of 1 L with d.H_2O.

CBS agar and CBS NS agar are a modification of CBS and CBS NS minimal media respectively for producing agar plates. 14 g Oxoid Agar Bacteriological (Agar No.1) is added to the 800 mL CBS or CBS NS part A solution pre-autoclave, then combined with part B and other growth supplements at ~55°C under sterile conditions. 20 mL of this agar solution were then added to sterile Petri dishes to produce CBS or CBS NS agar plates, and stored in the dark at 4°C.

2.2.3 Media supplements

As carbon-free and carbon/sulfur-free minimal media respectively, CBS and CBS NS media were supplemented with various carbon and/or sulfur sources for utilisation by *Hyphomicrobium* strains and species. These included methanol (MeOH), sodium sulfate (Na_2SO_4), dimethylsulfide (DMS), dimethylsulfoxide (DMSO) and dimethylsulfone ($DMSO_2$), methanesulfonic acid (MSA) and sodium methanesulfonate (NaMSA). Working stoks of each of these solutions were sterilised and added to CBS or CBS NS media post autoclave.

To produce working stocks of DMS and MeOH, 100% liquid phase stock was added to 20-100 mL deionised water in 120 mL serum vials that had been sterilised by autoclaving at 121°C for 15 min. Working stocks of DMSO and MSA were created by diluting the 100% stock solutions and with deionised water, then filter sterilised by passing them through a 0.2 µm Sartorius™ Minisart™ NML Syringe Filters. Working stocks of $DMSO_2$, DMSO, NaMSA and Na_2SO_4 were by dissolving the powder form of each compound in deionised water, which were then sterilised by autoclaving at 121°C for 15 min. The above compounds used as growth substrates were obtained from Sigma Aldrich or Thermofischer Scientific and

had the following purities: $DMSO_2$ ≥98.0%, DMSO ≥99.9%, DMS ≥99.0%, MeOH ≥99.9%, MSA ≥99.9%, NaMSA ≥98.0% and Na_2SO_4 ≥98.0%.

2.3 Cultivation of *Hyphomicrobium* species

Maintenance of *Hyphomicrobium* species was performed by cultivation on CBS or CBS NS agar plates or serum vials of CBS or CBS NS liquid media. Glycerol stocks of each species were streaked onto CBS agar plates and incubated at 30°C for 4-8 days in airtight gas jars containing gaseous MeOH. This was produced by adding 0.5 mL MeOH to a piece of cotton wool into an uncapped 5 mL universal and placing it at the bottom of a sealed gas jar to allow the MeOH to volatilise.

Single cell *Hyphomicrobium* colonies were then used to inoculate 120 mL serum vials containing 20 mL liquid cultures of CBS media with 20 mM MeOH, or CBS media with 20 mM MeOH and 1 mM Na_2SO_4. Serum vials were made airtight by fitting a rubber seal over the aperture, held in place by an aluminium crimp top cap. Liquid *Hyphomicrobium* cultures were incubated at 30°C with 150 rpm agitation for 24-72 h, then stored at RT for up to one month. Cultures were resubbed by the addition of 1 mL turbid culture to 20 mL of fresh media in 120 mL serum vials as needed.

Long term storage of *Hyphomicrobium* species was performed by cultivating each strain on liquid media in serum vials as described above for 24-48 h, then combined with 100% glycerol to produce 1 mL 40% glycerol stocks of *Hyphomicrobium* and stored at -80 °C. Note that the 100% glycerol used for producing glycerol stocks was sterilised by autoclaving at 121°C for 15 min prior to use.

Unless otherwise stated, cultivation of *Hyphomicrobium* species for experimentation was performed in 20 mL, 50 mL or 500 mL cultures with a headspace of approximately 100 mL, 200 mL and 1.5 L respectively. 20 mL cultures were grown using 20 mL of CBS or CBS NS media in airtight 120 mL serum vials, sealed with a rubber seal and aluminium crimp top cap. 50 mL cultures were grown in 50 mL CBS or CBS NS media in an airtight 250 mL conical flask sealed with a rubber cap. 500 mL cultures were grown in 500 mL CBS or CBS NS media in an airtight 2 L conical flask sealed with a rubber cap.

2.4 Measurement of bacterial cell density

Bacterial growth was monitored by measuring the optical density (OD) of bacterial cell cultures using a Biochrom Ultrospec 3100 *pro* spectrophotometer at a wavelength of 540 nm using a path length of 1 cm, in 1.0 mL plastic cuvettes. Unless otherwise stated, culture samples for spectroscopy were measured against a blank reference of appropriate media.

2.5 Phenotyping of *Hyphomicrobium* species

The phenotyping of MSC utilisation by *Hyphomicrobium* spp. was performed by cultivating individual *Hyphomicrobium* strains *H. denitrificans* ATCC51888, *H. facile* Bras3, *H. methylovorum* Bras1, *H. sulfonivorans* S1 and *H.* sp. VS, using MSCs as either sole carbon source or sole sulfur source, against positive (methanol) and negative (no carbon/ sulfur) controls.

2.5.1 MSCs as a carbon source

Hyphomicrobium spp. were cultivated in triplicate on the sole carbon sources of DMS, DMSO, $DMSO_2$ and MSA against a positive control of MeOH and a negative control without carbon. This involved growing a starter culture with a sole carbon source of MeOH, used to inoculate carbon-limited phenotyping cultures that were supplemented with additional carbon source over time and monitored for bacterial growth by measuring OD_{540} by spectroscopy.

Starter cultures of each *Hyphomicrobium* strain were cultivated in 20 mL CBS NS media in 120 mL serum vials, with a sole carbon source of 20 mM MeOH and sole sulfur source of 1mM Na_2SO_4. After 48 h growth at 30°C with 150 rpm of agitation, 0.5 mL of starter culture was used to inoculate fed-batch phenotyping cultures containing 20 mL CBS NS media with 1 mM Na_2SO_4 as inorganic sulfur source and a sole carbon source of either 1 mM MeOH for positive control cultures, no carbon for negative control cultures or 1 mM carbon for MSC cultures. This equated to 1 mM of MSA or 0.5 mM of DMS, DMSO or $DMSO_2$.

The cell density of inoculated culture at time zero was measured by removing a 1 mL sample and measuring optical density at 540 nm (OD_{540}), and the cultures incubated at 30°C with 150 rpm of agitation. At each consecutive sampling point 1 mL of sample was removed from the phenotyping culture for OD_{540} measurement and then replaced by loading the culture with 1 mL of fresh CBS NS media with 1 mM Na_2SO_4 and 20x concentration of initial culture carbon. This was either 20 mM of MeOH for positive control cultures, 10 mM MSC for dimethyl MSC cultures (DMS, DMSO and $DMSO_2$), 20mM MSA for MSA cultures or no carbon source for negative control cultures. carbon for MSC cultures, adjusted by molecular

carbon content of each carbon source. allowing for a gradual increase in carbon content over time while maintaining the same volume of the phenotyping culture.

2.5.2 MSCs as a sulfur source

Hyphomicrobium spp. were cultivated in triplicate on the sole sulfur source of DMS, DMSO, $DMSO_2$ and MSA against a positive control of Na_2SO_4 and a negative control without sulfur. This involved growing a starter culture with a sole sulfur source of Na_2SO_4, used to inoculate sulfur-limited phenotyping cultures that were monitored for bacterial growth by spectroscopy.

Starter cultures of each *Hyphomicrobium* strain were cultivated in 20 mL CBS NS media in 120 mL serum vials, with a sole carbon source of 20 mM MeOH and sole sulfur source of CBS NS media. After 48 h growth at 30°C with 150 rpm of agitation, 0.5 mL of starter culture was used to inoculate phenotyping cultures containing 20 mL CBS NS media with 1 mM Na_2SO_4 as inorganic sulfur source and a sole carbon source of either 0.1 mM Na_2SO_4 for positive control cultures, no sulfur for negative control cultures or 0,1 mM sulfur for MSC cultures. This equated to 0.1 mM of DMS, DMSO, $DMSO_2$ or MSA.

The cell density of inoculated culture at time zero was measured by removing a 1 mL sample and OD_{540} measured, and the cultures incubated at 30°C with 150 rpm of agitation. At each consecutive sampling point 1 mL of liquid was removed from the phenotyping culture for further OD_{540} measurement.

2.6 Genomics of *Hyphomicrobium* species

2.6.1 Sequencing of *Hyphomicrobium* genomes

Hyphomicrobium species for genome sequencing were cultivated on MeOH CBS agar plates to produce a lawn of cells. The biomass was then physically removed from the plate surface and submitted to the MicrobesNG genome sequencing service at the University of Birmingham.

MicrobesNG then performed DNA extraction from *Hyphomicrobium* biomass, followed by genome sequencing on an Illumina MiSEQ platform using 2x250 bp paired-end reads, which were trimmed using Trimmomatic (Bolger *et al.*, 2014) and combined using Bedtools (Quinlan and Hall, 2010). MicrobesNG then returned trimmed Illumina forward paired-end reads, reverse paired-end reads, forward unpaired reads and reverse unpaired reads for each *Hyphomicrobium* species.

2.6.2 Assembly of *Hyphomicrobium* genomes

Genome assembly of *Hyphomicrobium* species was performed using SPADES version 3.11.1 (Bankevich *et al*., 2012), an assembly tool for short genomes.

The assembly *H. sulfonivorans* S1 used a hybrid assembly that combined the new Illumina read data generated by MicrobesNG with a previously assembled contig of the *dmoA* gene cluster (NCBI Accession GQ980036) and an existing draft genome of *H. sulfonivorans* S1 (GOLD Project Gp0008840). SPADES' paired-end library function was used to incorporate the trimmed Illumina forward paired-end reads, reverse paired-end reads, forward unpaired reads and reverse unpaired reads into a single assembly. The *dmoA* cluster contig and draft genome of *H. sulfonivorans* S1 were used for the SPADES assembly as a high confidence 'trusted contig' and a low confidence 'untrusted contig' scaffold for assembling the MicrobesNG Illumina reads respectively. Finally, the SPADES 'careful' function was used to reduce the number of indels and mismatched reads following SPADES' recommended guidelines for genome assembly. This produced an output genome assembly for *H. sulfonivorans* in FASTA file format. Contigs smaller than 1kb in length were then removed from this new *H. sulfonivorans* S1 genome assembly to leave a total of 14 contigs, which were then reordered using the Mauve Contig Mover included in the Mauve Aligner (Rissman *et al*., 2009). This involved an initial rearrangement of the old *H. sulfonivorans* S1 draft genome performed against the new genome assembly, followed by an arrangement of the new assembly against the rearranged, old *H. sulfonivorans* draft genome.

```
spades.py --pe1-1 PAIRED_READS_1.fastq --pe1-2 PAIRED_READS_2.fastq --pe1-s
UNPAIRED_READS.fastq --careful --trusted-contigs TRUSTED_CONTIG.fasta --untrusted-contigs
UNTRUSTED_REFERENCE_GENOME.fa -o OUTPUT_DIRECTORY
```

Figure 2.7-1: *Command line input for genome assembly of* H. sulfonivorans *S1 using the SPADES platform.*

The genome assembly for *Hyphomicrobium methylovorum* Bras1 was performed using the 2x250bp Illumina paired and unpaired read data generated by the MicrobesNG with existing 2x80bp Illumina paired end read data previously generated by the Schäfer lab group (unpublished), to producing a more comprehensive assembly that would be possible with either single data set. SPADES' paired-end library function was used to incorporate the trimmed forward and reverse paired-end reads, unpaired reads reverse paired-end reads, forward unpaired reads and reverse unpaired reads into a single assembly.

```
spades.py --pe1-1 PAIRED_READS_1-1.fastq --pe1-2 PAIRED_READS_1-2.fastq --pe1-s
UNPAIRED_READS_1.fastq --pe2-1 PAIRED_READS_2-1.fastq --pe2-2 PAIRED_READS_2-
1.fastq --careful -o OUTPUT_DIRECTORY
```

Figure 2.7-2: *Command line input for genome assembly of* H. methylovorum *Bras1 using the SPADES platform.*

2.6.3 Annotation of *Hyphomicrobium* genomes

Annotation of *Hyphomicrobium* genomes was performed using the NCBI Prokaryotic Genome Annotation Pipeline as a Whole Genome Shotgun project (Angiuoli *et al.*, 2008). This was supplemented by a functional annotation of each genome using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa *et al.*, 2015), assigning a KO number (a unique KEGG reference for enzymes catalysing a specific chemical process) for each feature predicted by the NCBI genome annotation. This was performed via KEGG's BlastKOALA and BlastGHOST annotation tools (Kanehisa *et al*., 2016).

2.6.4 BLAST tools

Basic Local Alignment Search Tool (BLAST) searches were performed using the NCBI BLAST algorithm (Altschul *et al.*, 1990) via either the NCBI's internet browser based BLAST tools, or locally using the BLAST+ executables (Camacho *et al*., 2009). Browser-based BLAST searches were performed against NCBI's Nucleotide collection (nr/nt) or Non-redundant protein sequences (nr). BLAST+ searches were performed on an individual basis against specific genomes and/or metagenomic databases.

2.6.5 Phylogenetic tools

Phylogenetic analysis was performed using the Phylogeny.fr phylogenetics pipeline (Dereeper *et al.*, 2008): multiple amino acid sequence alignments were performed using MUSCLE version 3.8.31 (Edgar, 2004); maximum likelihood phylogenetic trees (bootstrap 100) were constructed using PhyML version 3.0 (Guindon *et al.*, 2010); phylogenetic trees were plotted using TreeDyn (Chevenet *et al*., 2006).

2.7 Batch culture of *Hyphomicrobium* species for comparative proteomics

Each proteomics experiment required the cultivation of *Hyphomicrobium* species to produce biomass for comparative proteomics. The specifications of these batch cultures are outlined below and summarised in Table 2.8-1.

Table 2.8-1: *Summary of substrate conditions for the proteomics of* Hyphomicrobium *species*

Strain	Condition	initial inoculum	
		1° carbon source	1° sulfur source
H. sulfonivorans S1	Control	20 mM MeOH	1 mM Na2SO4
	Carbon	10 mM $DMSO_2$	1 mM Na2SO4
	Sulfur	20 mM MeOH	1 mM $DMSO_2$
	Carbon/Sulfur	10 mM $DMSO_2$	
H. denitrificans ATCC 51888	Control	2 mM MeOH	1 mM Na2SO4
	DMS	1 mM DMS	1 mM Na2SO4
H. methylovorum Bras1	Control	2 mM MeOH	1 mM Na2SO4
	$DMSO_2$	1 mM $DMSO_2$	1 mM Na2SO4
	DMS	1 mM DMS	1 mM Na2SO4

2.8.1 Batch cultivation of *Hyphomicrobium sulfonivorans* utilising MSCs as carbon source

Initial starter culture of 20 mL CBS NS minimal media containing 20 mM MeOH and 1 mM Na_2SO_4, incubated at 30°C with 150 rpm agitation for 48 h. 1 mL aliquots of turbid starter culture were then used to inoculate the 1st round of control and treatment conditions in triplicate 20 mL cultures, containing CBS NS minimal media with an inorganic sulfur source of 1 mM Na_2SO_4 and sole carbon source of either 20 mM MeOH (control) or 10 mM $DMSO_2$ (treatment).

The two sets of 1st round triplicate cultures were incubated at 30°C with 150 rpm agitation until they reached an OD_{540} of ~0.150-0.250, occurring at 12 h and 24 h for the control and treatment conditions, respectively, at which point a 1 mL sample was removed and diluted

to an OD_{540} of 0.015 using CBS NS media. 1 mL of the adjusted aliquots from both conditions were then used to inoculate their respective 2nd round of control and treatment condition triplicate 500 mL cultures, again containing CBS NS media with sulfur source of 1 mM Na_2SO_4 and sole carbon source of either 20 mM MeOH (control) or 10 mM $DMSO_2$ (treatment).

The two sets of 2nd round triplicate cultures were incubated at 30°C with 150 rpm agitation until they reached an OD_{540} of ~0.150-0.250 representing *H. sulfonivorans* culture in exponential growth, occurring at 30 h and 36 h for the control and treatment conditions respectively. The *H. sulfonivorans* biomass from each culture was extracted and concentrated to ~50 mL by centrifugation for 15 min at 8670 g, 4°C, then further concentrated to 5 mL by centrifugation for 10 min at 4500 g, 4°C. The triplicate biomass samples for control and treatment conditions were then flash frozen in liquid nitrogen and stored at -80°C.

2.8.2 Batch cultivation of *Hyphomicrobium sulfonivorans* utilising MSCs as sulfur source

Initial starter culture of 20 mL CBS NS minimal media containing 20 mM MeOH and 1 mM Na_2SO_4, were incubated at 30°C with 150 rpm agitation for 48 h until turbid. 1 mL aliquots of turbid starter culture were then used to inoculate the 1st round of control and treatment conditions in triplicate 20 mL CBS NS minimal media cultures, namely control, sulfur treatment and carbon/sulfur treatment. Control treatment cultures contained a sole carbon source of 20 mM MeOH and an inorganic sulfur source of 1 mM Na_2SO_4, sulfur treatment cultures contained 20 mM MeOH and a sole sulfur source of 1 mM $DMSO_2$, and the carbon/sulfur treatment cultures contained a sole carbon and sulfur source of 20 mM $DMSO_2$.

The 1st round cultures for each condition were incubated at 30°C with 150 rpm for 24 h for the control and treatment conditions respectively, at which point a 1 mL sample was removed and diluted to an OD_{540} of 0.152 using CBS NS media. 0.5 mL of the adjusted aliquots from both conditions were then used to inoculate their respective 2nd round of control, sulfur treatment and carbon/sulfur treatment conditions in triplicate 500 mL cultures.

The three sets of 2nd round triplicate cultures were incubated at 30°C with 150 rpm for 30 h representing *H. sulfonivorans* culture in exponential growth, occurring at 30 h and 36 h for the control and treatment conditions respectively. The *H. sulfonivorans* biomass from each culture was extracted and concentrated to ~50 mL by centrifugation for 15 min at 8670 g, 4°C, then further concentrated to 5 mL by centrifugation for 10 min at 4500 g, 4°C. The triplicate

biomass samples for control and treatment conditions were then flash frozen in liquid nitrogen and stored at -80°C.

2.8.3 Batch cultivation of other *Hyphomicrobium* species utilising MSCs as carbon source

Hyphomicrobium spp. were cultivated in triplicate on the sole carbon sources of DMS and/or $DMSO_2$ against a positive control of MeOH. This involved growing a starter culture with a sole carbon source of MeOH, used to inoculate carbon limited phenotyping cultures that were monitored for bacterial growth by spectroscopy and supplemented with additional carbon source over time. *H. denitrificans* was cultivated using DMS as a treatment condition against a control of MeOH, while *H. methylovorum* using two treatment conditions, DMS and $DMSO_2$, against a control of MeOH.

Initial starter cultures of 20 mL CBS NS minimal media were produced containing 20 mM MeOH and 1 mM Na_2SO_4, were incubated at 30°C with 150 rpm agitation for 48 h until turbid. 1 mL aliquots of turbid starter culture was then used to inoculate 500 mL control and treatment cultures in triplicate, containing 500 mL CBS NS supplemented with an inorganic sulfur source of Na_2SO_4 and a sole carbon source of 2 mM carbon (either 2 mM MeOH, 1 mM DMS or 1 mM $DMSO_2$). Control and treatment cultures were then incubated at 30°C with 150 rpm agitation, measuring cell density at OD_{540} by extracting 1 mL culture volume. Cultures were further supplemented over the course of the experiment by the regular addition of 1 mL carbon source solution, consisting of CBS NS media with 1 mM Na_2SO_4 and sole carbon source of 1 M carbon (1 M MeOH, 0.5 M DMS or 0.5 M $DMSO_2$).

2.8 Comparative proteomics of *Hyphomicrobium* species

The comparative proteomics of *Hyphomicrobium* species involved the extraction and purification of protein from *Hyphomicrobium* biomass samples, an in-gel protein digestion to yield purified peptides and the submission of these peptides to the University of Warwick's Proteomics facility for mass spectrometry. These samples were then analysed using a comparative proteomics pipeline.

2.8.1 Protein extraction

Concentrated *Hyphomicrobium* biomass samples were thawed on ice and lysed by French press, in aliquots of 3-4 mL with a cell pressure of 2850 PSI. Samples were then centrifuged for 30 s at 13,000 g, 4°C to pellet cell debris and unlysed cells, and then transferred to a fresh Eppendorf tube. Protein samples were diluted in NuPAGE™ 4x LDS Sample Buffer (ThermoFisher Scientific) and denatured by incubation at 71°C for 10 min to disrupt secondary and tertiary protein structure, then placed back on ice for 10 min.

Evaluation of protein denaturation was performed by running each sample on an SDS-PAGE gel against a broad range protein standard (NEB Colour Protein Standard, Broad Range; New England Biolabs, UK), based on the observation of distinct protein banding. 5 µL of protein standard and 5-20 µL of each protein sample, equalised by sample culture cell density, were loaded onto Expedeon 4-12% SDS-PAGE pre-cast gels (Expedeon Ltd., Harston, UK) and run for 12 min at 65 V, followed by 200 V for 2 h. The SPS-PAGE gels were then treated with Expedeon InstantBlue™ Coomassie stain for 1 hr, washed with water and observed on a light box.

Purification of protein sample for mass spectroscopy was performed by running them through an SDS-PAGE gel in preparation for in-gel protein digestion. 20-40 µL denatured protein samples, adjusted to represent equal amounts based on sample culture cell density, were loaded onto Expedeon 4-12% SDS-PAGE pre-cast gels, leaving a lane between each sample to minimise risk of cross-contamination between protein samples. Gel electrophoresis was performed at 65 V for 12 min, followed by 200 V for 30-60 min, to a lane run length of ~2-3 cm. The gels were then treated with Expedeon InstantBlue™ Coomassie stain for 1 h, washed with water and observed on a light box. The protein containing lane of each sample was then excised from the gel, the top and bottom of the excised lane trimmed by ~1 mm and stored at -20°C for in-gel digestion.

2.8.2 In-gel protein digestion

SDS-PAGE gel fragments containing cellular proteins were washed, digested to generate peptides for submission to University of Warwick's Proteomics facility for mass spectroscopy, according to the following the protocol provided by the facility. This involved the sequential treatment of gel fragments from each sample with various enzymes and reagents to remove Coomassie stain, break down secondary protein structure, digest proteins and finally free the resultant peptides from the SDS page gel.

Protein-containing gel fragments from each biomass sample were loaded into 1.5 mL Eppendorf tubes and sliced into small ~2-4 mM cubes to increase their surface area. Fragments were then submerged in a 50 mM ammonium bicarbonate (ABC) solution in 50% (v/v) ethanol and incubated at 55°C for 20 min with 650 rpm and the liquid discarded to remove the gel's Instant-Blue Coomassie stain. This was repeated 3-4 times until the gel fragments were devoid of blue stain.

De-stained gel fragments were dehydrated by submersion of the gel in 100% ethanol for 5min with 650 rpm agitation, liquid removed by pipette and the dehydrated fragments submerged in a 10 mM dithiothreitol 10 mM ABC solution to mediate the reduction of protein disulfide bonds. After incubation at 56°C for 30 min with 650 rpm agitation, this solution was removed and replaced with a 55 mM 2-Iodoacetamide 50 mM ABC solution for the alkylation of cysteine residues, preventing the formation of new disulfide bonds.

The submerged gel fragments were incubated in the dark for 20 min at RT, liquid removed by pipette and the fragments washed three times by submersion in a solution of 50 mM ABC in 50% ethanol, incubation for 20 min at RT with 650 rpm agitation before the removal of the post-wash solution. Washed gel fragments were dehydrated once more by submersion of the gel in 100% ethanol for 5 min with 650 rpm agitation and the resultant liquid removed by pipette.

An in-gel digestion was performed on the gel fragments by the addition of 40 µL 50 mM ABC 2.5 ng/ µL bovine trypsin solution to rehydrate the gel, with 60 µL 50 mM ABC added to submerge the gel fragments as needed. The trypsin solution now permeating the gel fragments was allowed to digest the proteins suspended in the gel by incubating the fragments overnight at 37°C, and the liquid removed from each sample by pipette.

Peptide extraction from the post-trypsin digest gel fragments performed by suspension of the fragments 100 µL of a 25% acetonitrile 5% formic acid buffer to produce 100 µL peptide

sample solution. The peptide sample solution was transferred to a fresh Eppendorf tube and the peptide extraction step repeated twice more to produce 300 µL of peptide sample solution for each sample of *Hyphomicrobium* biomass.

300 µL peptide samples were then concentrated to a final volume of ~20 µL at 40°C for 3-6 h using a SpeedVac™ sample concentrator. Samples were then loaded with 10-30 µL of 2.5% acetonitrile 0.05% trifluoroacetic acid and submitted to the University of Warwick Proteomics department for mass spectroscopy. However, in the event that a protein sample contained a chemical contaminant, indicated by colouration of the concentrated peptide sample solution, peptides were loaded onto a C18 spin column for washing (see below).

2.8.3 C18 spin column peptide purification

A C18 filter column was produced for each sample by loading a 200 µL pipette tip with two layers of C18 membrane from Empore™ SPE Disks (3M, UK), then washed by loading 50 µL 100% methanol and centrifuged at 4000 g for 2 min. The column membrane was then equilibrated by loading 50 µL of 100% acetonitrile onto the column, centrifuged at 4000 g for 2 min, then loading 50 µL of 2% acetonitrile and 0.05% trifluoroacetic acid and centrifuged at 4000 g for 4 min. The flow-through was then discarded and the column ready for sample purification.

Each ~20 µL peptide sample was loaded onto an equilibrated C18 filter column and centrifuged 4000 g for 10 min. The flow-through was discarded and the column washed by the addition of 50 µL 2% acetonitrile 0.05% trifluoroacetic acid and centrifuged at 4000 g for 4 min. Finally, the column was transferred to a fresh Eppendorf tube and the washed peptide samples eluted from the column by loading it with 20 µL 60% acetonitrile followed by centrifuging at 4000 g.

The peptide containing flow-through for each sample was then concentrated to a final volume of ~10 µL at 40°C for 1-2 h using a SpeedVac™ sample concentrator, evaporating the acetonitrile. Concentrated samples were then combined with 10-40 µL of 2.5% acetonitrile 0.05% trifluoroacetic acid to a final volume of 50 µL and submitted to the University of Warwick Proteomics department for mass spectrometry.

2.8.4 Mass spectrometry

Mass spectrometry of peptides for proteomics was performed by the University of Warwick Proteomics facility. The following is a description of the mass spectrometry protocol provided by the facility:

The peptides were analysed with two columns, an Acclaim PepMap µ-precolumn cartridge 300 µm i.d. x 5 mM length, 5 µm particle size, 100 Å pore size and an Acclaim PepMap RSLC 75 µm i.d. x 25 cm, 2 µm, 100 Å (Thermo Scientific). The columns were installed on an Ultimate 3000 RSLCnano system (Dionex) at 40°C. Mobile phase buffer A was composed of 0.1% formic acid and mobile phase B was composed of acetonitrile containing 0.1% formic acid. Peptides were loaded onto the µ-precolumn equilibrated in 2% aqueous acetonitrile containing 0.1% trifluoroacetic acid for 8 min at 10 µL min-1 after which peptides were eluted onto the analytical column at 300 nL min-1 by increasing the mobile phase B concentration from 8% B to 25% over 78 min, then increased to 35% over 10 min, followed by a 5 min wash at 90% B and a 10 min re-equilibration at 4% B.

Eluting peptides were converted to gas-phase ions by means of electrospray ionization and analysed on a Thermo Orbitrap Fusion (Thermo Scientific). Survey scans of peptide precursors from 375 to 1500 m/z were performed at 120K resolution (at 200 m/z) with a 5x105 ion count target. The maximum injection time was set to 150 ms. Tandem MS was performed by isolation at 1.2 Th using the quadrupole, HCD fragmentation with normalized collision energy of 33%, and rapid scan MS analysis in the ion trap. The MS2 ion count target was set to 4x103 and maximum injection time was 200 ms. Precursors with charge state 2–6 were selected and sampled for MS2. The dynamic exclusion duration was set to 40 s with a 10 ppm tolerance around the selected precursor and its isotopes. Monoisotopic precursor selection was turned on and instrument was run in top speed mode.

2.8.5 Proteomics data analysis pipeline

Raw mass spectrometry data was searched against the proteomics sample's respective reference genome and the MaxQuant common contaminant database using MaxQuant version 1.5.5.1 (Tyanova et al., 2016a). Peptides were generated from the reference genome based on the predicted products of a trypsin digest with up to two missed cleavages, cysteine carbamidomethylation as fixed modifications, methionine oxidations as variable modifications and N-terminal acetylation and as a variable modification. Precursor mass tolerance was 10 ppm and product ions were searched at 0.8 Da tolerances.

Analysis of the processed data (LFQ-intensity) was performed using Perseus version 1.5.5.3 (Tyanova *et al.*, 2016b). Protein list was filtered to remove potential contaminants, reverse and proteins only identified by site to reduce risk of false positives. Proteins were considered to be positively identified if they were identified by at least one unique peptide in the triplicate samples of at least one condition, i.e. a protein would be discarded if it was

identified in one or two replicates of three different samples but kept if it was identified in all three replicates of the same sample set. If a protein was positively identified in at least one sample set, then the missing values of other samples would be replaced from their respective normal distributions. Note that although this may have given rise to an arbitrarily high fold-change between samples for some proteins (>500 fold-change) it is considerably more informative for a comparative analysis than giving these proteins a nominal abundance of zero (∞ fold-change). Statistical analysis was performed using a Student's T-test (p-value) and an adjusted Welch's T-test with a false discovery rate (FDR) of 0.05% (q-value). Samples were then counted as statistically significant if their q-value was less than 0.05 (q-value <0.05).

2.9 Continuous culture of *Hyphomicrobium sulfonivorans*

Cultivation of *Hyphomicrobium sulfonivorans* biomass for the comparative transcriptomics of MSC utilisation was performed in continuous culture to reduce the indirect effect of MSC utilisation when compared to a non-MSC carbon source. This was performed using FerMac 310/60 Bioreactor Fermenters (Electrolab, UK), ran in parallel on alternating growth substrates across several runs to generate pairs of 'control' and 'treatment' condition biomass samples cultivated on a sole carbon source of MeOH and $DMSO_2$ respectively.

Each chemostat run consisted of 6 stages: A pre-autoclave phase (P1) and post-autoclave phase (P2) to prepare the system for batch culture, batch culture phase (P3) to generate sufficient biomass to sustain continuous culture, a pre-continuous culture phase (P4) to prepare the system for continuous culture and finally a continuous culture phase (P5) to generate *H. sulfonivorans* biomass for chemostat biomass extraction (P6). The intention was to maintain stable, continuous culture for three volume changes prior to sampling for culture media to produce sufficient biomass for successful RNA extraction and submission for RNA-sequencing.

2.9.1 Chemostat technical specifications

The principle components of the FerMac 310/60 Bioreactor Fermenters used to cultivate *H. sulfonivorans* were a chemostat vessel, FerMac 310 Stirrer (FM310) and Fermac 360 Controller (FM360), that together modulated the culture's agitation, airflow, dissolved oxygen content, pH and temperature. The chemostat vessel of each chemostat was supplemented by an alkali stock for pH regulation, media vessel containing fresh media for continuous culture, waste vessel for culture effluent and sampling device for extracting culture

biomass. The identical specifications of two FerMac 310/60 Bioreactor Fermenters, henceforth called the system, are as follows:

Culture dilution rate was managed by using an external peristaltic pump to control flow of fresh media from the media vessel to the chemostat vessel, passing through an airlock to prevent contamination of the media vessel from the chemostat culture. Culture volume was limited to 1.8 L through overflow of excess culture via a pipe in chemostat vessel wall, leading to the waste vessel. The culture's dilution rate was calibrated in the pre-continuous culture phase by measuring the elution rate of the chemostat system.

The temperature of the vessel was measured by a temperature probe calibrated with a single point calibration at 30°C, connected to the FM360 and inserted into a port in the chemostat culture. Temperature of the culture was maintained at 30°C by the antagonistic activity of a cooling finger against a heat jacket, with the cooling finger receiving a flow of cold water from a water service line connected to the FM310 and inserted into culture, while the heat jacket was connected to the FM360 and fitted to the outside of the vessel.

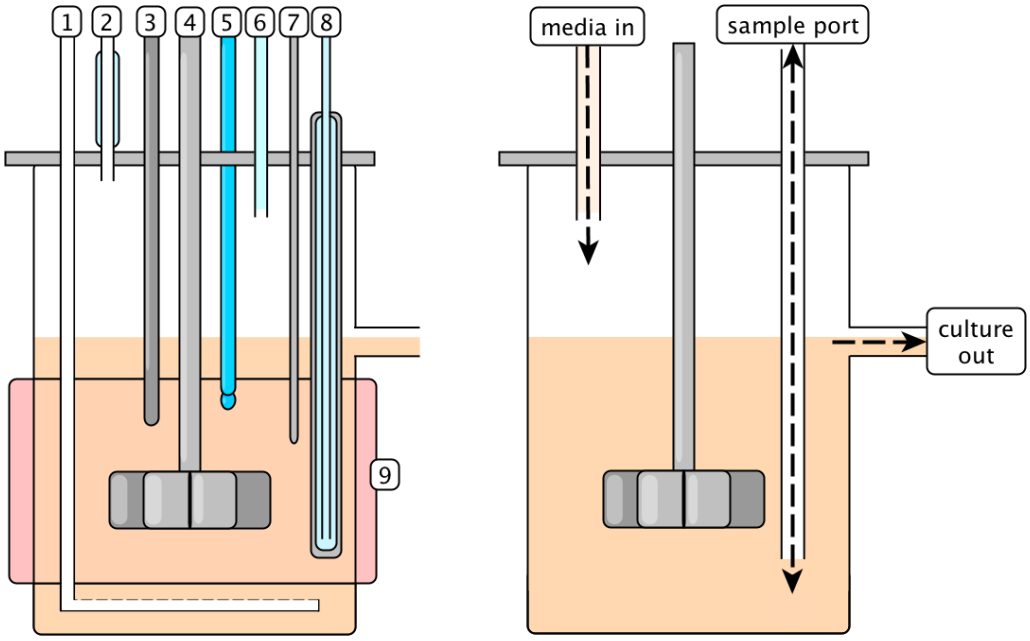


Figure 2.9-1: *Chemostat system schematics*. (LEFT) Representation of the chemostat control elements, including air intake with sparger (1), air outflow with condenser (2), dissolved oxygen electrode (3), impeller (4), pH electrode (5), alkali in-take (6), temperature probe (7), cooling finger (8) and heat jacket (9). (RIGHT) Representation of the chemostat flow system, including media intake for fresh media, culture outflow for effluent, and sampling port for inoculation and sampling.

Culture pH was monitored by a pH probe (Broadley-James Ltd, UK) connected to the FM360, attached to the chemostat vessel and inserted into the culture. Culture pH was maintained at pH 7.38 (±0.02) by an inflow of 1 M NaOH from the alkali stock into the chemostat vessel by a peristaltic pump controlled by the FM360. The pH probe was calibrated pre-autoclave with a two-point calibration at pH 7 and 10, then adjusted post-autoclave by measuring the pH of an aliquot of media taken from the autoclaved vessel. This was performed using an external pH probe with a three-point calibration at pH 4, 7 and 10.

Dissolved oxygen (DO) concentration was managed by variable agitation and constant aeration, monitored by a DO probe (Broadley-James Ltd, UK) inserted into the culture and connected to the FM360. Agitation of the culture was performed by an impeller inserted into the culture, connected to the FM310 and adjusted by the FM360 between a range of 150-350 rpm to regulate the culture DO. Aeration of the culture was provided by a constant air flow of 0.5 L/min from the FM310, passed through a 0.2 μm poresize filter and entering into the chemostat vessel via a sparger inserted into the culture below the impeller. The DO probe was calibrated in the post-autoclave phase using a single point calibration of the system at an arbitrary maximum DO of 100%, set after equilibration of the vessel for 15 min with 0.5L/min airflow, 350 rpm agitation and a temperature of 30°C.

Airflow out of the chemostat vessel primarily occurred through a water-cooled condenser with 0.2 μm poresize air filter, with additional 0.2 μm poresize filtered air outlets fitted to the alkali stock, media vessel, media airlock, sampling device and waste vessel.

2.9.2 Chemostat P1: Pre-autoclave

In the pre-autoclave phase, the pH and temperature probes were calibrated according to manufacturer's recommendation, followed by the assembly of the chemostat vessel with its acid/base inlet, condenser, cooling finger, DO probe, impeller, pH probe, media inlet, sampling port, sparger and temperature probe finger. The chemostat vessel and media vessels were then loaded with 1 L and 8 L of CBS NS media respectively. The chemostat vessel, media airlock, media vessel, sampling device and waste vessel were then sterilised by autoclave at 121°C for 15 minutes.

2.9.3 Chemostat P2: Post-autoclave

The post-autoclave phase involved the assembly of the chemostat system for batch culture by connection of the chemostat vessel to the FM310 module (condenser, cooling finger, impeller and sparger), 360 module (DO probe, heat jacket and pH probe) and sampling device

(sampling port). The chemostat vessel was loaded with 250 mL of CBS NS, air inflow was set to 0.5 L/min and the DO probe was calibrated. The chemostat vessel was then loaded with 12.5 mL of 2 M carbon source via the sampling port (either 2 M MeOH or 1 M $DMSO_2$) and 1.25 mL 1 M sulfur source to yield a final concentration of 1 mM Na_2SO_4 and 20 mM carbon (20 mM MeOH or 10 mM $DMSO_2$).

2.9.4 Chemostat P3: Batch culture

At the start of the batch culture phase, inoculation of the chemostat vessel was performed by loading 5 mL of OD_{540} 0.5-0.6 culture *H. sulfonivorans* S1 starter culture into the chemostat vessel. A single starter culture was used for each paired chemostat run, produced by cultivating *H. sulfonivorans* S1, in 20 mL CBS NS media with 20 mM MeOH and 1 mM Na_2SO_4, at 30°C for 48 h. Chemostat batch culture was agitated at 150 rpm, incubated at a constant temperature of 30°C (±0.2°C) and received an air flow of 0.5 L/min via the sparger. Sampling occurred at 0, 6, 24 and 48 h for OD measurement at 540nm by removal of 10 mL batch culture.

2.9.5 Chemostat P4: Pre-continuous culture

After 48 h of batch phase growth the system began the pre-continuous culture phase, in which the system was prepared for continuous culture. The autoclaved media vessel containing 8 L CBS NS Media part A, was loaded with 2 L of CBS NS media part B (phosphate buffer solution), carbon source to 20 mM carbon and 1 mM sulfur source. The media vessel was then connected to the media pump and airlock, followed by connection of the chemostat vessel to the airlock and waste container. The media pump was then activated to allow media to flow into the chemostat vessel at 180 mL/hr until full, ready for continuous culture.

2.9.7 Chemostat P5: Continuous culture

In the continuous culture phase the chemostat system was set to cultivate *H. sulfonivorans* at a constant cell density, dilution rate, DO content, pH and temperature. Dissolved oxygen content of the chemostat culture was maintained at 30-40% by an air flow of 0.5 L/min into the culture and varying the culture agitation from 150-350 rpm to change culture aeration. Culture pH was maintained at pH 7.37 (±0.03) by addition of 1 M NaOH to counteract acidification of *H. sulfonivorans* culture over time. The temperature of the culture was maintained at 30°C (±0.2°C).

Culture dilution rate was maintained at 180 mL/h of CBS NS media with 1 mM $NaSO_4$ and a sole carbon source of either 20 mM MeOH for the control condition or 10 mM $DMSO_2$

for the test condition, an equal quantity of carbon for both conditions. This 180 mL/h dilution rate translated to a volume change of 10% total culture per hour, requiring 30h of uninterrupted continuous culture for the three volume changes of media chosen for this experiment.

2.9.8 Chemostat P6: Biomass extraction and RNA preservation

Generation of biomass for RNA sequencing was performed over several weeks, so to prevent RNA degradation biomass was treated with the RNA preservative Invitrogen™ RNA*later*™ Stabilization Solution (Invitrogen, Thermofisher Scientific, UK), flash frozen in liquid nitrogen and stored at -80°C.

For each sample over 200 mL of continuous culture was extracted from its respective chemostat and split into 50 mL aliquots. This was immediately centrifuged at 4500 g for 10 min at 4°C to pellet the biomass and the culture supernatant discarded.

RNA preservation was performed by suspending the four cell pellet aliquots in 0.5-1.0 mL RNA*later* (1 mL per 0. 8 OD_{450} of bacterial culture), incubating them at RT for 10 min and then 2 h at 4°C to allow to RNA*later* to permeate cells. The time from the extraction of bacterial biomass to the addition of RNA*later* was kept to less than 30 min to limit risk of sample degradation. Samples were then flash frozen in liquid nitrogen and stored at -80°C until RNA extraction.

2.10 Comparative transcriptomics of *Hyphomicrobium sulfonivorans*

Preserved RNA from *H. sulfonivorans* biomass was purified using the spin-column based Qiagen RNeasy® Mini Kit and submitted to Novogene (Novogene Ltd., Hong Kong) for RNA sequencing on an Illumina HiSEQ™ to produce paired-end read data for comparative transcriptomics. This data was then aligned to the *H. sulfonivorans* genome, read counts quantified for each feature as predicted from the organism's annotated genome and then examined using a differential gene analysis.

2.11.1 RNA extraction protocol

RNA extraction was performed for each sample using the Qiagen RNeasy® Mini Kit, adapted from the manufacturer's protocol "Purification of total RNA from bacteria using the RNeasy® Mini Kit" and "RNeasy Mini Handbook". All centrifugation was performed in a benchtop centrifuge at 4°C.

For each sample, one 1 mL aliquot of frozen RNA*later* suspended cells was thawed on ice and pelleted at 13,000 g for 2 min, the supernatant removed and the cells resuspended in 350 µL of Qiagen RLT lysis buffer to mediate chemical cell lysis. 10 µL 14.3 M β-mercaptoethanol was then added to each sample to mediate denaturation of RNAses.

The cell suspension was loaded into 2 mL lysis matrix tubes containing 0.1 mm silica beads and physically lysed by ball milling three times, in which samples were incubated on ice for 2 min and homogenised at 6 m/s for 30 s on a FastPrep-24 sample homogeniser (MP Biomedicals, UK). The resultant lysate was then separated from cell debris and silica beads by centrifugation at 13,000 g for 15 s and the lysate transferred to a fresh tube.

Sample lysate was combined with an equal volume of 70% nuclease-free ethanol and loaded onto a Qiagen RNeasy spin column, then centrifuged at 13,000 g for 15 s to remove waste liquid. The column membrane containing the bound RNA was then washed by loading 350 µL Buffer RW1 to the column followed by centrifugation at 13,000 g for 15 s, the flow-through was discarded.

To denature DNA from the lysate loaded onto the column, 80 µL of Qiagen DNAse I solution was then loaded onto to the spin column membrane followed by incubation at RT for 15 min. This was then washed from the column by loading 350 µL Buffer RW1 to the column membrane and centrifugation performed at 13,000 g for 15 s, with flow-through discarded. Two further wash steps were then performed by loading 500 µL of Buffer RPE to the column, followed by centrifugation at 13,000 g and 4°C for 15 s in the first wash, and 2 min in the second wash.

To elute RNA, the column was transferred into a new tube and loaded with 50 µL nuclease-free water, then centrifuged at 13,000 G for 1 min. To increase the yield of RNA from the column, the eluent was then loaded back onto the column and centrifuged again at 13,000 g for 1 min. Samples were then stored at -20°C.

Quantification of RNA samples was performed using a Qubit® RNA HS Assay Kit, then submitted for qualitative analysis by the University of Warwick Genomics Facility using a Bioanalyser with the RNA Nano chip.

2.11.2 RNA sequencing

H. sulfonivorans RNA samples were submitted to Novogene (UK) Company Limited for ribosomal RNA depletion, library preparation and Illumina RNA sequencing. Enrichment of messenger RNA by ribosomal RNA depletion was performed using the Illumina Ribo-Zero

rRNA Removal Kit (Bacteria). Library prep performed using the NEB Next® Ultra™ RNA Library Prep Kit. Illumina sequencing was performed at a depth of 6 million reads using an Illumina HiSEQ™, with 150 bp paired-end reads.

2.11.3 Transcriptomics data analysis pipeline

The forward and reverse paired-end reads for each *H. sulfonivorans* biomass sample, generated by RNA-sequencing were aligned to the new *H. sulfonivorans* assembly, counted and compared to generate transcriptome data using the following pipeline:

Alignment of raw read data using Bowtie2

Alignment of paired-end Illumina reads to the *H. sulfonivorans* S1 draft genome 2018 (BioProject PRJNA437222) was performed using Bowtie2 version 2.3.4 (Langmead and Salzberg, 2012). The *H. sulfonivorans* genome in FASTA format was used to build a Bowtie reference genome using the 'bowtie2-build' function. The forward and reverse paired-end reads for each sample were then aligned against the Bowtie reference genome using the 'Bowtie' function, using the 'very-sensitive' preset for performing sensitive alignments. This produced an output alignment file in the 'sam' format for each sample.

```
>bowtie2-build REFERENCE_GENOME.fasta BOWTIE_REFERENCE_GENOME

>bowtie2 -x BOWTIE_REFERENCE_GENOME -1 PAIRED_READS_1.fastq -2 PAIRED_READS_2.fastq -S ALIGNED_READS.sam --fr --very-sensitive-local -p10
```

Figure 2.11.1: *Command line input for Bowtie2 raw read alignment to reference genome.*

SAM (Sequence Alignment/MAP) format RNA alignments to the *H. sulfonivorans* genome were then converted to the BAM (Binary Alignment/MAP) format using SAMtools (Li *et al.*, 2009), via the 'view' function. BAM format aligned reads were then sorted using the SAMtools 'sort' function to reorder the aligned sequences and indexed using the 'index' function to allow them to be viewed using the alignment viewers Artemis and Integrative Genomics Viewer (IGV).

```
>samtools view –bS ALIGNED_READS.sam > ALIGNED_READS.bam

>samtools sort ALIGNED_READS.bam SORTED_ALIGNED_READS

>samtools index SORTED_ALIGNED_READS.bam
```

Figure 2.11.2: *Command line input for SAMtools conversion of data from SAM to BAM format.*

Read counting using StringTie

The number of reads aligning to predicted coding sequences in the *H. sulfonivorans* S1 genome were then counted using StringTie version 1.3.4b, following the recommended workflow for differential expression analysis (Pertea *et al.*, 2016). For each sample, the aligned read data was assembled into transcripts based on the predicted coding sequences in the reference genome (GTF FORMAT). These mapped transcripts were then merged into a single reference list of potential isoforms as a pseudo reference genome in the GTF file format. The aligned read data was then assembled for each sample again, this time against the pseudo reference genome (GTF) to generate count data for each sample (GTF).

Differential expression analysis using DESeq2

The GTF count data files were converted into TXT files for differential expression analysis, then passed through DESeq2 version 1.18.1 (Anders and Huber, 2010), using R version 3.4.3. DESeq2 normalises the data, removes outliers and provides expression data in fold-change expression. A paired differential expression analysis was performed for the MeOH samples versus the samples cultivated on $DMSO_2$. Each MeOH and $DMSO_2$ sample was paired with their respective partner based on chemostat run, i.e. sample MeOH 1 from Run 1 was paired with $DMSO_2$ from Run 1.

```
>./stringtie SORTED_ALIGNED_READS.bam -B -C -e -G
ANNOTATED_REFERENCE_GENOME.gff -o FEATURE_READ_COUNTS.gtf

>./stringtie --merge -p 8 -G ANNOTATED_REFERENCE_GENOME.gff -o
PSEUDO_REFERENCE_GENOME.gtf LIST_OF_SAMPLES_TO_MERGE.txt

>./stringtie SORTED_ALIGNED_READS.bam -B -C -e -G
PSEUDO_REFERENCE_GENOME.gtf -o
FEATURE_AND_NOVEL_ISOFORM_READ_COUNTS.gtf
```

Figure 2.11.3: *Command line input for StringTie read counting.*

A DESeq2 sample table was produced detailing the experimental parameters, sample files and sample identities. The condition of each sample defined as ‘sample condition’ was assigned as either a ‘control’ condition for samples cultivated on MeOH or a ‘treatment’

condition for those cultivated on $DMSO_2$. The chemostat run for each sample was then defined as 'sample number', assigned for each sample as '1', '2' or '3' to enable normalisation between the three pairs of chemostat runs that had been used to generate RNA samples.

A DESeq2 database was constructed using the 'ddsHTSeq' function, then a differential expression analysis performed for the sample set using the 'dds' function. This was then processed into a data matrix using the 'res' function with an 'alpha' of "0.05", optimising independent filtering for adjusted p-values of ≤0.05 (Bourgon *et al*., 2010). The resultant data matrix contained the log2-fold change, p-value and adjusted p-value (padj) of each feature in the *H. sulfonivorans* genome based on a comparison of the control samples cultivated on MeOH and the treated samples cultivated on $DMSO_2$, normalised based on chemostat run.

```
>sampleFiles<-
c('MEOH_1.txt','MEOH_2.txt','MEOH_3.txt','DMSO2_1.txt','DMSO2_2.txt','DMSO2_3.txt')

>sampleCondition<-c('control','control','control','treated','treated','treated')

>sampleNumber<-c('1','2','3','1','2','3')

>sampleTable<-data.frame(sampleName=sampleFiles, fileName=sampleFiles,
condition=sampleCondition, number=sampleNumber)

>ddsHTSeq<-DESeqDataSetFromHTSeqCount(sampleTable=sampleTable,
directory="OUTPUT_FOLDER", design=~ number + condition)

>ddsHTSeq$condition

>dds<-DESeq(ddsHTSeq)

res <- results(dds, alpha=.05)
```

Figure 2.11.4: *R input for DESEQ2 analysis.*

2.11 Gas chromatography

Quantification of DMS in bacterial culture was performed by gas chromatography (GC) on a Shimadzu GC-2010 Plus (Shimadzu UK, Milton Keynes, UK), by comparing the concentration of gaseous DMS in culture headspace to standards of known concentration; DMS standards were equivalent to their respective cultures supplemented with an appropriate range of DMS concentrations to generate a five-point standard curve.

During GC analysis, 100 μL of headspace gas, from a either a culture or DMS standard was loaded onto the GC system – fitted with a Shim-1 capillary column (30.0 m x 0.5 mm), using helium as carrier gas, heated to 180^{o}C and connected to a flame photometric detector – to generate peak data for integration. Sample peak size was then compared back to the standard curve to quantify culture DMS concentration.

Chapter 3:

Genomics and phenotyping of MSC metabolism in *Hyphomicrobium sulfonivorans* S1

3. Genotyping and phenotyping of MSC metabolism in *Hyphomicrobium sulfonivorans*

3.1 *Hyphomicrobium sulfonivorans* S1, a model organism of MSC metabolism

Hyphomicrobium sulfonivorans strain S1 (sp. nov) is a facultative methylotroph capable of utilising the MSC dimethyl sulfone ($DMSO_2$) as a sole carbon source, that has been studied sporadically for almost two decades (Borodina *et al.*, 2000). This organism has been chosen as the principal model organism for this project due to its ease of cultivation, range of previous research carried out in the organism and the research questions surrounding MSC metabolism in the organism that remain unanswered. Below is a brief summary of the previous work performed on *H. sulfonivorans* strain S1 (type strain), followed by an overview of the research described in this chapter. Note that although another strain of *H. sulfonivorans* has also been discovered, WDL6 (Albers *et al*., 2018a, Albers *et al*., 2018b), the name *H. sulfonivorans* will be used exclusively to refer to the type strain of this organism, *H. sulfonivorans* strain S1, unless explicitly stated otherwise (Borodina *et al.*, 2000; Borodina *et al.*, 2002).

Isolation and characterisation

H. sulfonivorans type strain S1 was isolated from garden soil in Warwickshire (UK) in an enrichment experiment by Borodina *et al.* (2000), in which soil samples were suspended in minimal media and supplemented with a carbon source of $DMSO_2$. They characterised the organism as a Gram-negative, prosthecate, rosette-forming member of the *Hyphomicrobium* genus with a cell size of ~0.25µm (see Figure 3.1-1).

Further strain typing by Borodina *et al.* (2000) showed that *H. sulfonivorans* S1 could grow on the methylated sulfur compounds dimethylsulfone ($DSMO_2$), dimethylsulfoxide (DMSO) and dimethylsulfide (DMS) as a sole carbon source, as well as the other C1 compounds methanol (MeOH), formaldehyde (HCHO), formate (CHOOH), monomethylamine (MMA) and trimethylamine (TMA).

Oxygen electrode experiments and enzyme activity assays were performed by Borodina *et al.* (2000) on cell-free extracts of *H. sulfonivorans*, cultivated on $DMSO_2$ as a sole carbon source. This led to the identification of DMS, DMSO, $DMSO_2$ and MSA-dependent NADH-oxidation activity in the *H. sulfonivorans* lysate, which was attributed to DMS monooxygenase, DMSO reductase, $DMSO_2$ reductase and MSA monooxygenase respectively. Further NADH-oxidation assays discovered that this DMSO and $DMSO_2$-dependent NADH-oxidation was

~10-fold greater in the membrane fraction compared to the cytosolic fraction, suggesting that the DMSO and $DMSO_2$ reductases were membrane bound. Oxygen electrode experiments were also able to identify the respiration of methanethiol (MT), suggesting the presence of an NADH-independent MT oxidase.

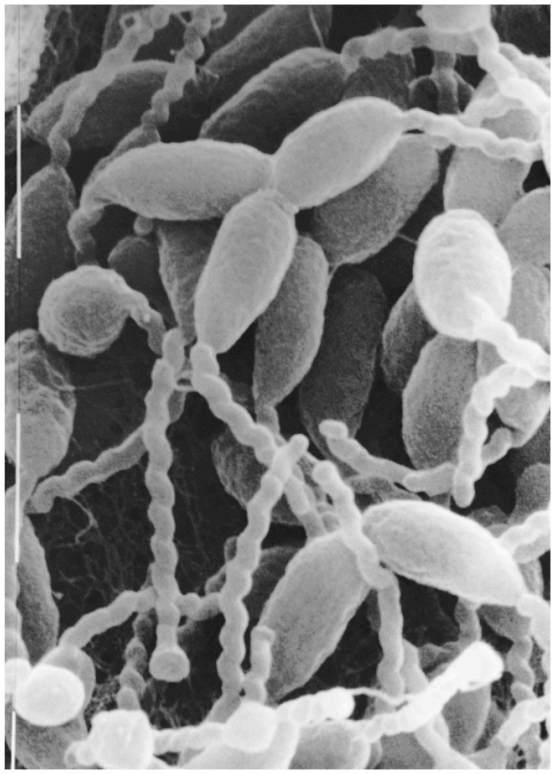


Figure 3.1-1: *Scanning electron micrograph of Hyphomicrobium sulfonivorans type strain S1* (Borodina *et al.*, 2000). Image distinctly shows the presence of helical prosthecates sprouting from the bacterial cell bodies, with multiple cell bodies attached together to form a rosette. The white scale bars on the left side of the image represent 1 μm.

This led Borodina *et al.* (2000) to propose the metabolic pathway of $DMSO_2$ utilisation in *H. sulfonivorans* strain S1 outlined in Figure 3.1-2, consisting of $DMSO_2$ reduction to DMS, followed by oxidation of DMS to hydrogen sulfide (Borodina *et al.,* 2000). In this process, $DMSO_2$ would be reduced to DMSO by a membrane-bound NADH-dependent $DMSO_2$ reductase, which would be further reduced to DMS by a membrane-bound NADH-dependent DMSO reductase. DMS would then be oxidised to MT and formaldehyde by an NADH-dependent DMS monooxygenase, with the MT being further oxidised to hydrogen sulfide and formaldehyde by an oxygen-dependent MT oxidase. It should be noted that a fifth enzyme, an

NADH-dependent MSA monooxygenase, was also proposed, but not included in this pathway due to the inability of *H. sulfonivorans* S1 to utilise the compound as a carbon source (Borodina *et al.* 2000).

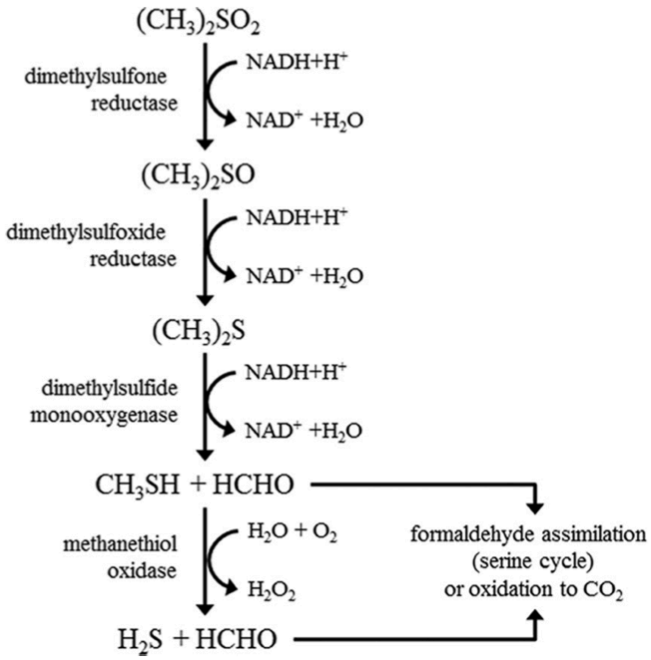


Figure 3.1-2: *Proposed pathway of $DMSO_2$ degradation in H. sulfonivorans* (Boden *et al.*, 2011a). Following the DMS monooxygenase pathway outlined in Figure 2, $DMSO_2$ is reduced by an NADH-dependent $DMSO_2$ reductase to DMSO, which is then further reduced to DMS by an NADH-dependent DMSO reductase (Borodina *et al*., 2002). An NADH-dependent DMS monooxygenase then oxidises DMS to form MT and formaldehyde (Boden *et al*., 2011), and an MT oxidase oxidises MT to produce formaldehyde, hydrogen peroxide and hydrogen sulfide.

Further research into MSC metabolism in *H. sulfonivorans* strain S1 performed by (Boden *et al.*, 2011a) led to the purification and characterisation of a two-subunit DMS monooxygenase, the first and only enzyme of MSC metabolism that has currently been identified in *H. sulfonivorans*. This heterodimeric enzyme, consisting of a large 53 kDa DmoA subunit and a small 19 kDa DmoB subunit, was found to be highly expressed when the organism was cultivated on $DMSO_2$ as a sole carbon source but notably absent when cultivated on MeOH, suggesting that it played an important role in MSC metabolism. Boden *et al.* (2010) used protein sequencing mass spectroscopy to establish an N-terminal sequence for the large 53 kDa subunit, identified the protein as a putative flavin-dependent monooxygenase based on sequence and performed enzyme activity assays on both subunits. It was found that the subunit mediated the NADH-dependent oxidation of DMS in a reaction that could be greatly improved

by the addition of FMN, allowing the subunit to be characterised as the $FMNH_2$-dependent monooxygenase DmoA of an NADH-dependent DMS monooxygenase.

Current research questions

The identification of the DmoAB enzyme by Boden *et al.* (2010) supported the pathway of $DMSO_2$ metabolism originally proposed by Borodina *et al.* (2000), in which $DMSO_2$ is reduced to DMS and then oxidised to hydrogen sulfide (see Figure 3.1-3), but substantial gaps in our knowledge still remain. Firstly, three of the four enzymes proposed for this pathway have yet to be characterised in *H. sulfonivorans*, namely the NADH-dependent membrane bound enzymes $DMSO_2$ reductase and DMSO reductase, and the NADH-independent enzyme MT oxidase. Furthermore, although the DmoA subunit of DMS monooxygenase was linked to the *dmoA* gene by Boden *et al.* (2011) they were unable to identify the gene encoding the 19kDa second subunit of the enzyme experimentally, predicted to be an NADH-dependent flavin oxidoreductase and prospectively called *dmoB*. Several candidates for this gene were found 3-6 kb upstream of the *dmoA* gene, but there is as yet insufficient evidence to assign identities to these candidates.

Finally, it is unknown whether $DMSO_2$ metabolism via DMS monooxygenase represents the sole mechanism for MSC metabolism in the organism or if there are alternate enzymes and pathways that can degrade these compounds for assimilation into *H. sulfonivorans* biomass. An example may be the unidentified MSA monooxygenase proposed by Borodina *et al.* (2000) as the source of MSA oxidation in *H. sulfonivorans* lysate, though it is unknown if this represents part of an alternate mechanism of $DMSO_2$ degradation or simply an unrelated or ancillary property of another enzyme system. Fortunately a disruption mutant of the DMS monooxygenase encoding *dmoA* gene, in which an internal sequence of *dmoA* was replaced by a gentamycin resistance cassette, was generated by Julie Scanlan of the Schäfer lab group prior to the start of this project (unpublished). This provides an opportunity for examining the role of the DMS monooxygenase and the associated DMS oxidation pathway in *H. sulfonivorans*.

Comparative and functional omics of MSC metabolism

The purpose of the research described in this chapter was to study *H. sulfonivorans* as a model organism of MSC metabolism, with the intention of better understanding the molecular mechanisms that mediate the degradation and assimilation of these compounds. The chapter's

work flow began by cultivating and phenotyping the *H. sulfonivorans* S1 wild-type (WT) strain against the DMS monooxygenase gene disruption mutant *ΔdmoA* (Section 3.2), to determine the range of MSCs that the organism can utilise as a carbon and sulfur source and establish the impact of *dmoA* disruption on MSC metabolism. Genome sequencing has been used to generate an improved reference genome (Section 3.3), which was used in functional genomics to search for putative enzymes of MSC metabolism and construct metabolic pathways (Section 3.4) for functional genomics in Chapter 4.

3.2 Phenotyping *Hyphomicrobium sulfonivorans* S1 strains

The first task of the project has been to confirm the range of MSCs that *H. sulfonivorans* S1 is able to utilise as a carbon and sulfur source. Although previous research by Borodina *et al.* (2000) indicated that the organism was able to use DMS, DMSO and $DMSO_2$ as a sole carbon source, the extent to which S1 can utilise MSCs has yet to be documented. The second task was to examine the hypothesis that:

The methylotrophic DMS monooxygenase pathway is the sole mechanism by which* H. sulfonivorans *can utilise MSCs.

This has been put forward based on the existing model of assimilatory MSC utilisation in *H. sulfonivorans* in which $DMSO_2$ is sequentially reduced to DMS, which is oxidised to sulfide to produce formaldehyde for methylotrophic carbon assimilation. It has then been proposed that this hydrogen sulfide might then be utilised as a sulfur source for assimilatory inorganic sulfur assimilation (Borodina *et al.*, 2000; Boden *et al.*, 2011).

To examine if the aforementioned hypothesis is true, the DMS monooxygenase gene disruption mutant *H. sulfonivorans* S1 *ΔdmoA* (*ΔdmoA*) has been phenotyped against its parent wild-type strain *H. sulfonivorans* S1 WT (WT) on a range of MSCs as carbon and sulfur sources. If the aforementioned hypothesis is true, then disrupting the *dmoA* gene encoding the DMS monooxygenase alpha subunit DmoA should prevent the organism from utilising $DMSO_2$ as a sole carbon source and, if also possible on the WT, as a sole sulfur source. However, if the hypothesis is false then we may expect to see MSC assimilation by *ΔdmoA* when utilising MSCs as a sole carbon and/or sole sulfur source.

3.2.1 Carbon Assimilation Phenotyping

H. sulfonivorans WT and *ΔdmoA* strains have been cultivated on range of MSCs as a sole carbon source to assess the impact of disrupting the *dmoA* gene on MSC metabolism and establish the range of compounds that *H. sulfonivorans* S1 is capable of utilising. Both strains were cultivated in triplicate on minimal media with a sole carbon source of either $DMSO_2$, DMSO, DMS or MSA against growth on MeOH as a control and a negative control condition without carbon. All cultures were supplemented with sulfate as a replete inorganic sulfur source. As DMS toxicity has previously been reported to exhibit bacterial growth at higher concentrations, it was decided that the MSCs would be introduced to *H. sulfonivorans* incrementally to induce growth.

Carbon-limited media was inoculated in parallel for each strain and condition in triplicate, then cultivated for 144 h with sampling of the cultures at 0, 48, 96, 120 and 144 h (see Figure 3.2-1). Growth assays were performed by measuring sample optical density (OD) at 540 nm. Cultures were then supplemented with additional carbon source in an equal volume of fresh culture, maintaining a consistent culture volume throughout the experiment. The WT showed significant growth on the positive control of MeOH and $DMSO_2$ as a sole carbon source, but not DMSO, DMS, MSA or the no carbon negative control. The *ΔdmoA* strain was also able to grow on MeOH and unable to grow on a sole carbon source of either DMSO, DMS, MSA or the no carbon negative control, but unlike the WT was unable to utilise $DMSO_2$ as a sole carbon source. The OD_{540}, growth yield and signifance for each strain/condition at T_{144} is displayed in Supplementary Table S3.2-1.

Further phenotyping of the *H. sulfonivorans* WT and *ΔdmoA* strains was performed to examine the degradation and/or generation of DMS by the organism when cultivated on MSCs. Both strains were cultivated in parallel in carbon limited batch cultures with a sole carbon source of MeOH for 120 h to generate biomass and deplete the carbon source within the cultures. Biomass was then then subdivided into smaller cultures and supplemented with either MeOH, $DMSO_2$, DMSO or DMS to a final concentration of 1 mM and the concentration of DMS in the culture headspace measured by gas chromatography at 1, 25 and 121 h. Although DMS was successfully detected in the *H. sulfonivorans* cultures supplemented with DMS there was neither a significant difference in DMS concentration between the two strains nor reduction in DMS concentration over the course of the experiment (see Supplementary Figure S3.2-1). Furthermore, no DMS was detected by gas chromatography in any of the cultures supplemented with other substrates, so if DMS generation occurred then it was below the limit of detection.

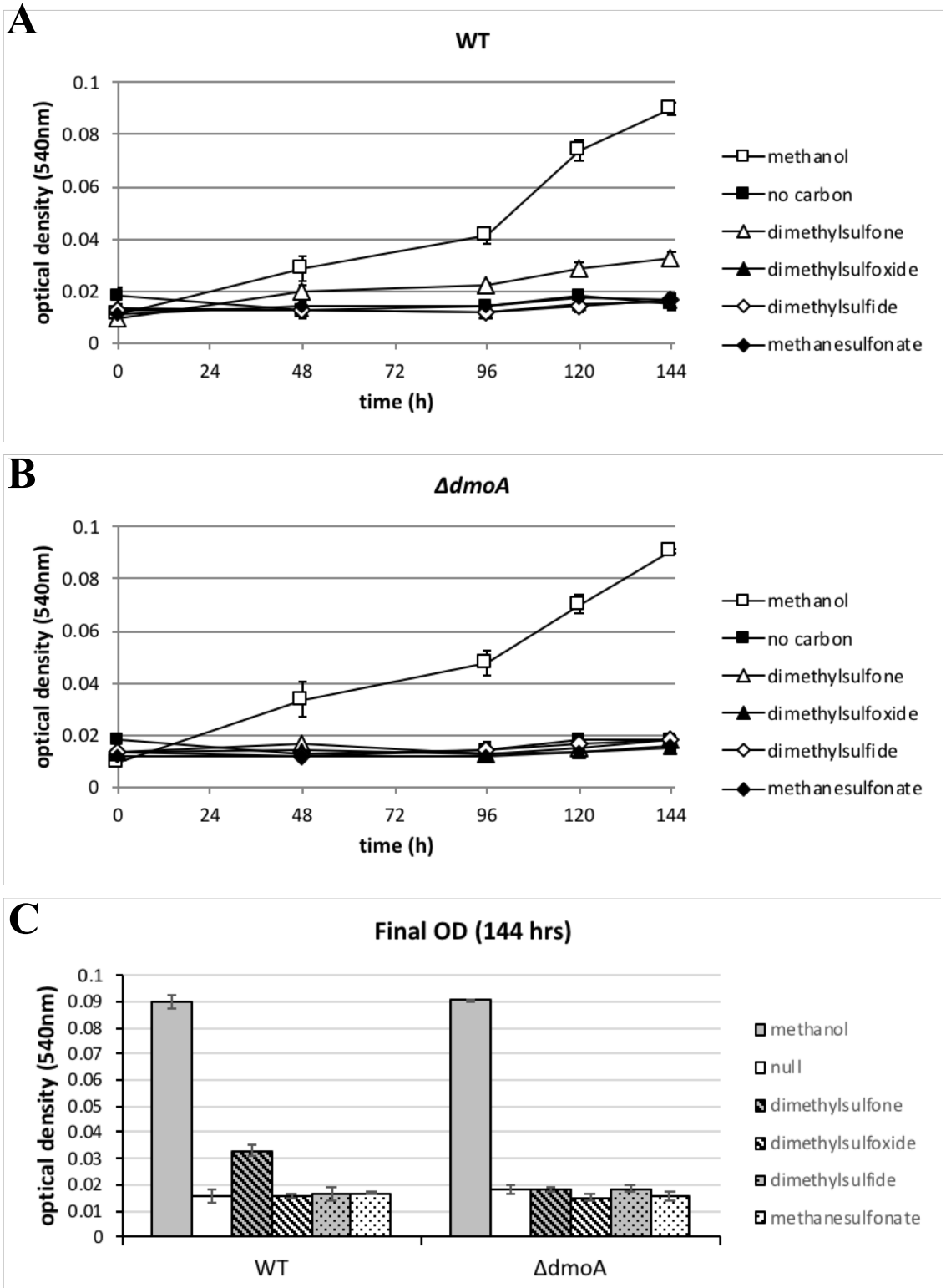


Figure 3.2-1: *H. sulfonivorans carbon source phenotyping.* Growth curves of the *H. sulfonivorans* (A) WT and (B) *ΔdmoA* strains cultivated on a range of MSCs as sole carbon source against a methanol positive control against a null carbon negative control, with the addition of more carbon at each sampling point. (C) Comparison of final OD_{540} for both strains at 144 h.

3.2.2 Sulfur Assimilation Phenotyping

The DMS monooxygenase pathway proposed by Boden *et al.* (2011) is thought to produce hydrogen sulfide as a product of DMS metabolism, making it possible that *H. sulfonivorans* could also utilise MSCs as a sole sulfur source. To determine which MSCs, if any, can be used by *H. sulfonivorans* as a sulfur source and to further examine the effect of *dmoA* deletion on the organism's MSCs utilisation, the WT and *ΔdmoA* strains were phenotyped for the utilisation of MSCs as a sulfur source. Both strains were cultivated in triplicate sets of cultures containing a carbon source of 20mM MeOH and supplemented with either $DMSO_2$, DMSO, DMS, MSA, sodium sulfate (positive control) or no sulfur source (negative control). Note that the carbon:sulfur ratio of the media varied between 100:1 (positive and negative control), 101:1 (MSA) and 102:1 ($DMSO_2$, DMSO and DMS), but that this 1-2% difference has been considered nominal for purpose of assessing net bacterial growth.

H. sulfonivorans cultures for each strain were inoculated in parallel for each condition in triplicate, then cultivated for 36 h with sampling of the cultures at 0, 12, 36 and 50 h (see Figure 3.2-2). The data indicates clear growth of *H. sulfonivorans* on a sole sulfur source of $DMSO_2$, DMSO and MSA that is comparable to the growth of the organism on the sulfate positive control, and modest growth of the organism on DMS that nevertheless exceeded that of the null sulfur negative control. Furthermore, no substantial change was seen in bacterial growth between the *H. sulfonivorans* WT and the gene disruption mutant *ΔdmoA*, suggesting that disrupting the *dmoA* gene does not impact the utilisation of MSCs as a sole sulfur source. The OD_{540}, growth yield and signifance for each strain/condition at T_{144} is displayed in Supplementary Table S3.2-2.

Further phenotyping of the *H. sulfonivorans* WT and *ΔdmoA* strains was performed to examine the degradation and/or generation of DMS by the organism when cultivated on MSCs. Both strains were cultivated in parallel in sulfur limited batch cultures with a sole carbon source of MeOH for 120 h to generate biomass and deplete the sulfur source within the cultures. Biomass was then then subdivided into smaller cultures and supplemented with either MeOH, $DMSO_2$, DMSO or DMS to a final concentration of 1 mM and the concentration of DMS in the culture headspace measured by gas chromatography at 0, 48 and 120 h.

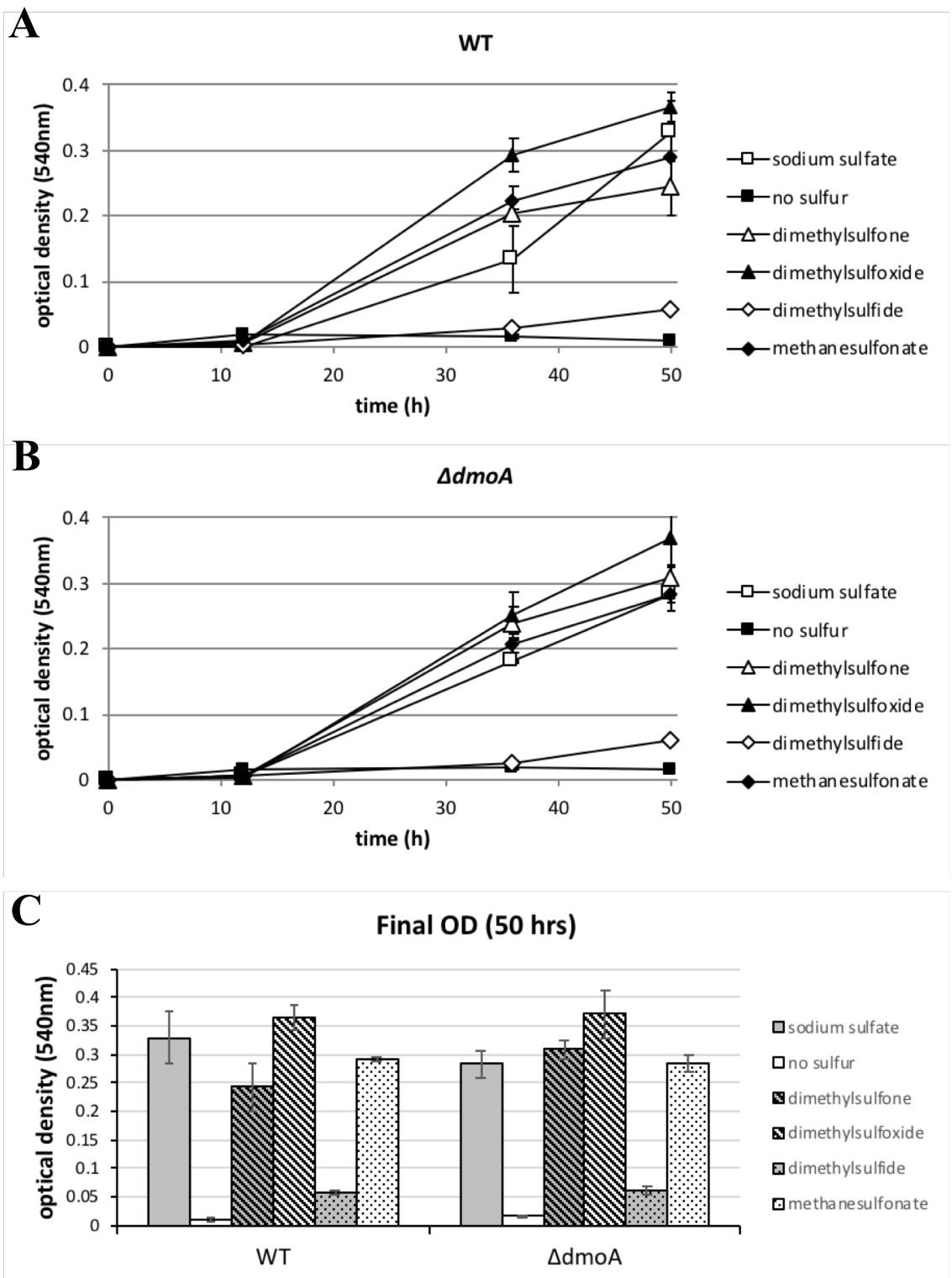


Figure 3.2-2: *H. sulfonivorans sulfur source phenotyping.* Growth curves of the *H. sulfonivorans* (A) WT and (B) *ΔdmoA* strains in batch cultivated on a range of MSCs as sole sulfur source against a sodium sulfate positive control and a null sulfur negative control. (C) Comparison of final OD_{540} for both strains at 50 h.

Although DMS was successfully detected in the *H. sulfonivorans* cultures supplemented with DMS there was neither a significant difference in DMS concentration between the two strains nor a reduction in DMS concentration over the course of the experiment (see Supplementary Figure S3.2-2). Furthermore, no DMS was detected by gas chromatography in any of the cultures supplemented with other substrates, so if DMS generation occurred then it was below the limit of detection.

3.2.3 Conclusions

The most significant finding from the phenotyping of these *Hyphomicrobium* strains is that although the disruption of the DMS monooxygenase encoding *dmoA* gene prevents the growth of the organism on $DMSO_2$ as a sole carbon source, the organism is still capable of growth on $DMSO_2$, DMSO, DMS and MSA as a sole sulfur source. This rejects the initial hypothesis that *H. sulfonivorans* S1 utilises MSCs exclusively via the proposed DMS oxidation pathway outlined by Boden *et al.* (2011).

What this suggests is that *H. sulfonivorans* has an alternate enzyme or metabolic pathway consisting of several enzymes that is capable of assimilatory MSC metabolism as a sulfur source, but not as a carbon source. Interestingly, it also suggests that this sulfur-specific pathway is capable of degrading a wider range of MSCs than the assimilatory DMS oxidation pathway under certain conditions.

An alternative metabolic pathway for *H. sulfonivorans* was briefly suggested by Borodina *et al.* (2000) in which $DMSO_2$ was sequentially oxidised to methanesulfinate (MSIA), MSA and then sulfite, based on the detection of MSA monooxgygenase activity in enzyme assays. Although they found that *H. sulfonivorans* could utilise $DMSO_2$, DMSO and DMS as a sole carbon source, the organism was unable to utilise MSA or oxidise $DMSO_2$ in an oxygen electrode experiment, pointing instead towards a pathway of $DMSO_2$ reduction and what would eventually be the DMS monooxygenase pathway (Boden *et al.*, 2011).

According to the initial growth experiments on *H. sulfonivorans* S1 by Borodina *et al.* (2000) following its isolation, the organism should be capable of growth on DMSO and DMS as a sole carbon source. However, numerous attempts throughout the project to cultivate the organism on these compounds have failed to promote sustained growth on these compounds (data not shown), despite the successful cultivation of several other *Hyphomicrobium* species on DMSO and DMS (see Chapter 5). This may suggest a mischaracterisation of bacterial growth for the organism in the original study, though the organism's inability to use DMSO or

DMS as a sole carbon source is puzzling given that these compounds are predicted to be intermediates of methylotrophic $DMSO_2$ metabolism.

Taken together, these results suggest the presence of two distinct mechanisms of MSC metabolism in *H. sulfonivorans*, a methylotrophic DMS oxidation pathway that involves a DMS monooxygenase and another unknown pathway that is incapable of utilising $DMSO_2$ as a sole carbon source, but nonetheless able to use other MSCs as a sole sulfur source. However, without also disrupting this second putative pathway of MSC utilisation it remains unknown if sulfur assimilation is possible via the DMS oxidation pathway.

3.3 Genome sequencing of *Hyphomicrobium sulfonivorans* S1

Performing an omics analysis of MSC metabolism in *H. sulfonivorans* required an accurate genome to use as a reference for genomics, proteomics and transcriptomics. Previous research into *H. sulfonivorans* strain S1 by Schäfer (unpublished) has generated a draft genome for the organism; Schäfer performed a hybrid assembly using 454 and Illumina sequence data (read length 70 bp) to generate *H. sulfonivorans* strain S1 2011 draft (GOLD Project Gp0008840). This draft genome has 3,986,593 base pairs split into 32 contigs, with 3,771 predicted coding sequences. The genome assembly also has a GC content of 61.2%, which is close to the 62% GC content measured by Borodina *et al.* (2000) during their earlier study of *H. sulfonivorans* strain S1. However, an inspection of the *dmoA* gene cluster by the Schäfer lab group revealed a disparity between the genome and an existing sequence for the region collected by Boden *et al.* (2011) and suggested a flaw in the genome which may be problematic for future omics analysis (Schäfer, unpublished).

It was therefore decided that the *H. sulfonivorans* genome would be re-sequenced in the hope of generating a more complete, more accurate genome to use as a reference for a multi-omics analysis. *H. sulfonivorans* strain S1 biomass was cultivated in monoculture and submitted to the MicrobesNG sequencing service for paired-end Illumina sequencing, performed on an Illumina MiSeq. This generated 150 bp paired-end raw reads that were then assembled into a new genome using the SPADES assembly platform (Bankevich *et al.*, 2012) in a hybrid assembly that assembled the new raw read data using two sets of the reference data: The S1 2011 draft genome assembled by the Schäfer lab group (GOLD Project Gp0008840) and the partial *dmoA* gene cluster (NCBI Accession GQ980036) sequenced by Boden *et al.* (2011).

This new genome for *H. sulfonivorans* strain S1 had a total of 3,766,243 base pairs split between 461 contigs, with 94.3% of base pairs located on contigs greater than or equal to 1 kb in length. The genome displayed an average GC content of 61.8% and average coverage depth ~26x. A visual representation of this information can be seen in Figure 3.3.1.

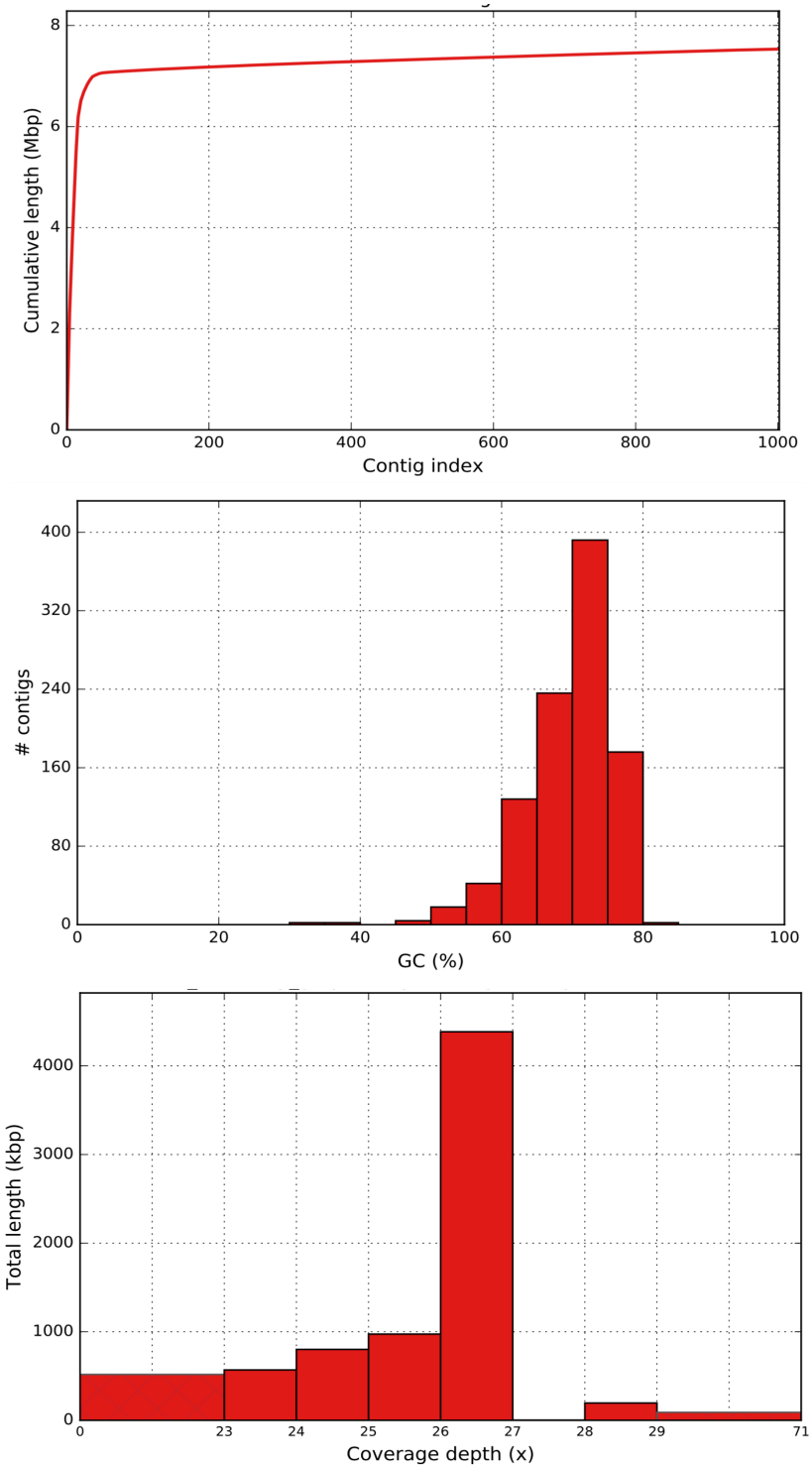


Figure 3.3-1: *Statistical plots of the* H. sulfonivorans *S1 2018 draft genome pre-removal of short contigs (<1000 bp).* (TOP) Cumulative length the genome represented by the minimum number of contigs. (MIDDLE) Histogram displaying the number of contigs of a given GC content (%) in brackets of 10%. (BOTTOM) Histogram displaying the average contig length (kbp) coverage depth within a given bracket of coverage depth (x). Statistical plots have been generated using the Quast genome quality assessment tool (Gurevich *et al.*, 2013). Note that nucleotide length is doubled due to assessment of forward and reverse sequences.

Contigs shorter than 1 kb were considered too small for multi-omics analysis and discarded from the genome, totalling 214,512 base pairs across 437 contigs. Although the loss of 5.7% of the total assembly is unfortunate it was decided to be a necessary compromise to prevent the misannotation of multi-omics data. This reduced genome had 3,551,731 bp spread over 14 contigs and has been used throughout the project for comparative genomics, proteomics and transcriptomics. The genome was then submitted to NCBI as a Whole Genome Shotgun (WGS) project for annotation using the Prokaryotic Genomes Annotation Pipeline (Haft *et al.*, 2017), under BioProject PRJNA437222 and Biosample SAMN08644077.

The GC content of the new *H. sulfonivorans* draft S1 2018 draft genome is 61.8%, marginally closer to the 62% found experimentally by Borodina *et al.* (2000) than the 61.2% found in the old assembly by the Schäfer lab group (GOLD Project Gp0008840). To compare the quality of the old and new *H. sulfonivorans* genome assemblies, a BLAST search was performed querying the *H. sulfonivorans* partial *dmoA* cluster previously sequenced Borodina *et al.* (2000) (NCBI Accession GQ980036) against both assemblies (see Figure 3.3-2). This *dmoA* cluster represents an important domain of MSC metabolism in the organism with respect to MSC metabolism (Boden *et al.*, 2011).

This generated a complete alignment of the 14,158 base pair partial *dmoA* cluster onto Contig 4 of the new draft genome, with a sequence identity of ~100% (14,156/ 14,158 bp), and a partial alignment to a 632 bp sequence of Contig 6 with only ~4.5% coverage (632/14,158 bp) but 100% sequence identity. Based on the NCBI annotation of the new genome it appears that this 632 bp sequence encodes a putative RibE riboflavin synthase alpha subunit, which is flanked by an MFS transporter and RibH-type synthase on Contig 4, but a RibD-type bifunctional deaminase/reductase and RibB-type synthase on Contig 6. It is therefore possible that this gene represents a genuine duplication of the *ribE* gene rather than a misassembly of the new genome.

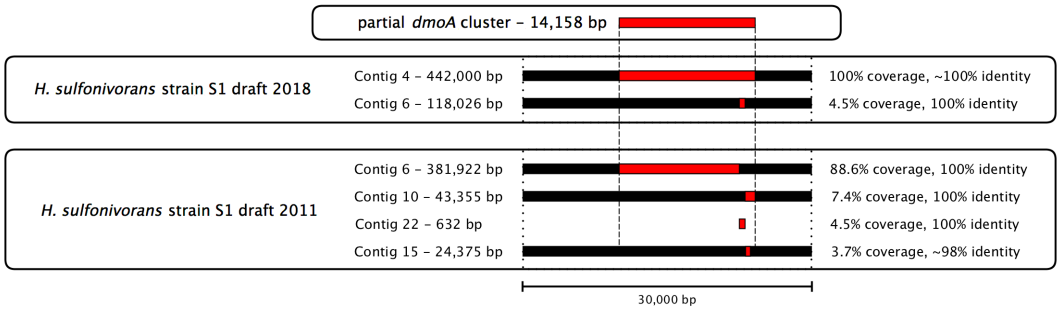


Figure 3.3-2: *BLAST search of the partial dmoA gene cluster against H. sulfonivorans strain S1 draft genomes*. Alignment of the partial *dmoA* gene cluster (NCBI Accession GQ980036) onto contigs of the two *H. sulfonivorans* strain S1 genomes, draft 2018 (this project) and draft 2011 (GOLD Project Gp0008840), with aligned sequences highlighted in RED.

In contrast, the alignment of the partial *dmoA* cluster to the S1 2011 draft generated a fragmented alignment with the largest fragment having a coverage of 88.57% (12,540/ 14,158 bp) and a sequence identity of 100%, which also featured an additional 632 bp encoding *ribE* sequence as an independent 632 bp contig. Of the remaining 1,618 bp sequence from the partial *dmoA* cluster, partial alignments were made on two different contigs in the old genome; one has a coverage of 64.6%1,046/ 1,618 bp) with a sequence identity of 100%, the other has a coverage of 32.4% (513/ 1,619 bp) and a sequence identity of 97.7% (513/ 525 bp).

A search for 16S RNA gene sequences in the *H. sulfonivorans* S1 2018 draft genome on the basis of NCBI and KEGG orthology (see Section 3.4 below) has identified two identical copies of the 16S RNA gene with the accession numbers C6Y62_05155 and C6Y62_05190. A BLASTn search for this 1,457 bp sequence against the NCBI nucleotide collection returns the 16S genes of 99% sequence identity (E-value 0.0) to *H. sulfonivorans* strains WDL6 (NCBI Accession AF538931.1), 25S (NCBI Accession AY305006.1), CT (NCBI Accession AY468372.1) and S1 (NCBI Accession NR_025082.1) – the latter being the 16S RNA gene sequence *H. sulfonivorans* S1 (sp. nov) submitted by Borodina *et al.* (2000). A complete copy of the 16S RNA gene sequence and its alignment to the S1 reference sequence is shown in Supplementary Figure 3.3-1.

Given that the new *H. sulfonivorans* S1 2018 draft genome resolves the multiple incomplete alignments of the partial *dmoA* cluster to the S1 2011 draft into one continuous sequence whilst also integrating the duplicate *ribE* sequence into a second *rib*-gene containing gene cluster, this suggests that the new 2018 assembly is a more accurate representation of the *H. sulfonivorans* genome, although it remains an incomplete draft. The *H. sulfonivorans* S1 2018 draft genome has been used as a reference genome for omics analysis in this organism.

3.4 Genome analysis of MSC metabolism in *Hyphomicrobium sulfonivorans*

The simplest way to identify the unknown enzymes of MSC metabolism in *H. sulfonivorans* may be to search the organism's genome for homologues of known enzymes that have already been identified and characterised in other organisms. This can also be expanded to include the assimilation and dissimilation of the proposed products of MSC metabolism, such as formaldehyde and inorganic sulfur. To this end, a combination of literature review, BLAST searches and gene annotation have been used to generate putative metabolic pathways of MSC metabolism in *H. sulfonivorans* for the mapping of proteomics and transcriptomics data. Note that all BLAST searches (tBLASTn, protein query versus nucleotide reference) were performed using the NCBI BLAST tool (Boratyn *et al.*, 2013), while gene annotations were performed using the NCBI Prokaryotic Genomes Annotation Pipeline (Haft *et al.*, 2017) and functional KEGG BlastKOALA annotation (Kanehisa *et al.,* 2016).

Based on what is already known about MSC metabolism in *H. sulfonivorans*, this can be split into three processes: MSC metabolism (Section 3.4.2), formaldehyde metabolism (Section 3.4.3) and inorganic sulfur compound metabolism (Section 3.4.4). The metabolism of MeOH is also an important process for understanding MSC metabolism, as it has been chosen as the methylotrophic control compound for the phenotyping and omics analysis of *Hyphomicrobium* species (Section 3.4.1). Note that as the *H. sulfonivorans* S1 genome used for this analysis is a draft, it may generate false negatives with respect to the presence of certain genes and errors with respect to identity and gene synteny. Each gene has received a locus tag consisting of a genome ID "C6Y62" and a unique gene ID "_00001", i.e. "C6Y62_00005".

3.4.1 Methanol Metabolism

Methanol (MeOH) metabolism in methylotrophs typically involves the oxidation of MeOH by a MeOH dehydrogenase to produce formaldehyde for subsequent metabolism and assimilation into the organism's biomass. The net sum of this reaction is that one molecule of MeOH produces one molecule of formaldehyde, two protons and two electrons. The simplicity of MeOH's oxidation metabolites compared to those of more complex C1 compounds such as methylated amines or methylated sulfur compounds makes MeOH an ideal control substrate for studying C1 compound metabolism using comparative omics. The results of a BLAST search for common enzymes of MeOH metabolism in the *H. sulfonivorans* genome are displayed in Table 3.4-1, while the successfully identified enzymes of MeOH metabolism have been mapped onto a putative metabolic pathway in Figure 3.4-1.

Two NADH-dependent MeOH dehydrogenases that have been characterised in the Alphaproteobacteria *Methylobacterium extorquens* are the calcium-containing MxaF-type MeOH dehydrogenase and lanthanide-containing XoxF-type MeOH dehydrogenase (Schmidt *et al.*, 2010, Nakagawa *et al*., 2012). The MxaF-type MeOH is a heterotetramer formed from two MxaF large subunits and two MxaI small subunits, encoded by the organism's *mxaFJIG* cluster (Amaratunga *et al.*, 1997). In addition to the large and small subunits, the cluster also encodes the cytochrome c(L) MxaG (Choi *et al.*, 2013) and another MxaJ, which mediates electron transfer between the MxaFI methanol dehydrogenase and MxaG (Amaratunga *et al.*, 1997).

A BLAST search for the large subunit MxaF shows one strong homologue in the *H. sulfonivorans* genome with 82% sequence identity, C6Y62_00385, supported by NCBI and KEGG annotations matching this enzyme (K14028). A search for small subunit MXaI points to C6Y62_11495 as the likely candidate, with 73% sequence identity and similarly supported by KEGG annotation (K14028). Further BLAST searches for MxaJ and a BLAST and KEGG orthology search for MxaG (K16255) identified the respective homologues C6Y62_00380 and C6Y62_00375, both with 54% sequence identity. Together, this suggests that all four genes of the *mxaFJIG* MeOH oxidation system appear to be present in *H. sulfonivorans* within the same gene cluster, albeit in the order of *mxaFJGI* (C6Y62_00385, C6Y62_00380, C6Y62_00375, C6Y62_00370).

The XoxF-type MeOH dehydrogenase is similar to the MxaF subunit of the MxaFJIG system, even sharing the system's small subunit MxaI (Schmidt *et al*., 2010). A BLAST search for this enzyme in *H. sulfonivorans* shows two homologues with just over 50% sequence identity, C6Y62_12520 and C6Y62_00385, the latter of which is now characterised as MxaF, making C6Y62_12520 the most likely *xoxF* candidate. However as XoxF-type enzymes represent a diverse and poorly characterised array of MeOH dehydrogenases that can be subdivided into at least five clades from XoxFI to XoxF5 (Chistoserdova, 2011; Keltjens, 2014), it may be premature to ascribe this relatively weak homologue a more precise function. Indeed, a BLASTx search for C6Y62_12520 homologues against the UniProt database (Apweiler *et al*., 2004) was unable to identify any experimentally characterised reference sequences for the predicted protein product of this gene.

The search for homologues of MxaF and XoxF has also highlighted a third putative MeOH dehydrogenase, C6Y62_12300, showing weak homology of 32% or 34% to MxaF and XoxF of *Methylobacterium extorquens* respectively. A reciprocal BLAST search against the UniProt database (Apweiler *et al*., 2004) suggests that it may be an alcohol dehydrogenase,

based on the strong homology of C6Y62_12300 to the quinoprotein alcohol dehydrogenase ExaA of *Pseudomonas aeruginosa* associated with ethanol oxidation (Görisch, 2003). This has been supported by a reciprocal BLAST search of this enzyme against the *H. sulfonivorans* genome, indicating a sequence identity of 68% for C6Y62_12300.

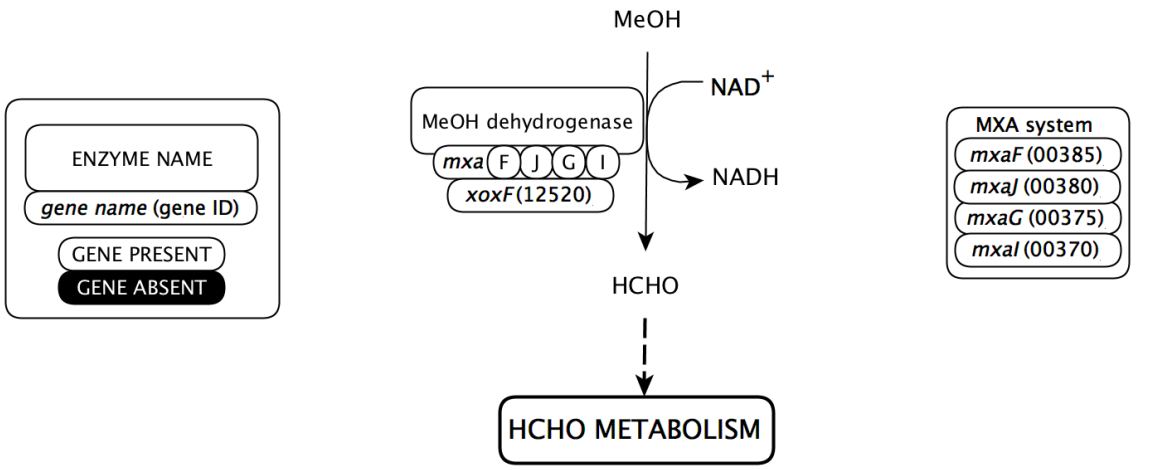


Figure 3.4-1: *Comparative genomics analysis of MeOH metabolism in* H. sulfonivorans. General legend for metabolic pathways: Enzymes with homologues in the *H. sulfonivorans* genome have been annotated with the gene name and gene ID. Genes with no homologue in the *H. sulfonivorans* genome are labelled BLACK.

Table 3.4-1: *BLAST search of the* H. sulfonivorans *strain S1 2018 draft for genes of methanol metabolism. Proteins highlighted in* ***bold*** *have been selected as strong homologues.*

METHANOL METABOLISM									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Mxa system MeOH dehydrogenase	subunit 1	MxaF	UNIPROT/P16027	*Methylobacterium extorquens*	**C6Y62_00385**	**PQQ-dependent dehydrogenase, methanol/ethanol family**	**82%**	**1018**	**0.0**
					C6Y62_12520	PQQ-dependent dehydrogenase, methanol/ethanol family	37%	363	2e-116
					C6Y62_12300	PQQ-dependent dehydrogenase, methanol/ethanol family	32%	228	2e-66
					C6Y62_11495	pyrrolo-quinoline quinone	29%	198	9e-56
	subunit 2	MxaI	UNIPROT/P14775	*Methylobacterium extorquens*	**C6Y62_00370**	**methanol dehydrogenase**	**73%**	**112**	**6e-34**
Mxa system protein		MxaJ	UNIPROT/P16028	*Methylobacterium extorquens*	**C6Y62_00380**	**methanol oxidation system protein MoxJ**	**54%**	**283**	**1e-94**
Mxa system cytochrome c-L		MxaG	UNIPROT/P14774	*Methylobacterium extorquens*	**C6Y62_00375**	**cytochrome c(L) MauG**	**54%**	**174**	**5e-55**
					C6Y62_00435	cytochrome c, class I	39%	83.2	7e-21
					C6Y62_00200	cytochrome-c oxidase	31%	38.9	9e-05
MeOH dehydrogenase		XoxF	UNIPROT/C5ATJ3	*Methylobacterium extorquens*	**C6Y62_12520**	**PQQ-dependent dehydrogenase, methanol/ethanol family**	**50%**	**580**	**0.0**
					C6Y62_00385	PQQ-dependent dehydrogenase, methanol/ethanol family	51%	536	0.0
					C6Y62_12300	PQQ-dependent dehydrogenase, methanol/ethanol family	34%	259	8e-78
					C6Y62_11495	pyrrolo-quinoline quinone	29%	211	2e-60
quinoprotein alcohol dehydrogenase		ExaA	UNIPROT/Q9Z4J7	*Pseudomonas aeruginosa*	C6Y62_12300	PQQ-dependent dehydrogenase, methanol/ethanol family	68%	837	0.0
					C6Y62_00385	PQQ-dependent dehydrogenase, methanol/ethanol family	36%	283	7e-87
					C6Y62_11495	pyrrolo-quinoline quinone	33%	263	3e-79
					C6Y62_12520	PQQ-dependent dehydrogenase, methanol/ethanol family	33%	235	1e-68

3.4.2 Methylated Sulfur Compound Metabolism

The metabolism of MSCs to generate precursors for biomass assimilation, such as formaldehyde and inorganic sulfur, can be split into three distinct categories: (i) the reversible oxidation and reduction reactions between dimethyl sulfur compounds (DMSCs), (ii) a DMS oxidation pathway that generates formaldehyde and hydrogen sulfide, and (iii) a $DMSO_2$ oxidation pathway that generates formaldehyde and sulfite (Schäfer *et al*., 2009). The results of a BLAST search for common enzymes of methylotrophic MSC metabolism in the *H. sulfonivorans* genome are displayed in Tables 3.4-2 (a, b), while the successfully identified enzymes of MSC metabolism have been mapped onto a putative metabolic pathway in Figure 3.4-2.

Interconversion of dimethyl sulfur compounds

The study of MSC metabolism in *Hyphomicrobium* species has found experimental evidence for an NADH-dependent $DMSO_2$ reductase, reducing $DMSO_2$ to DMSO, and an NADH-dependent DMSO reductase, reducing DMSO to DMS (Borodina *et al.*, 2000).

Although neither of these enzymes have been successfully characterised in any *Hyphomicrobium* species, the DMSO reductase DmsABC has previously been identified in *E. coli* (Sambasivarao *et al*., 1990). This is an NADH-dependent, trimeric enzyme consisting of the subunits DmsA, DmsB and DmsC, which catalyses several reduction reactions including DMSO to DMS, and trimethylamine N-oxide (TMAO) to trimethylamine (TMA). However, as a BLAST search against the *H. sulfonivorans* genome returns only poor homologues of DmsA and DmsB (<30% sequence identity), it makes DmsABC an unlikely candidate for a DMSO reductase in the organism.

Another potential mechanism of MSC metabolism in *H. sulfonivorans* is the sequential oxidation of DMS to DMSO by some form of DMS dehydrogenase, and DMSO to $DMSO_2$ by a DMSO dehydrogenase. A known example of a DMS dehydrogenase has been characterised in the purple phototrophic bacteria *Rhodovulum sulfidophilum* and consists of the subunits DdhA, DdhB and DdhC (McDevitt *et al.*, 2002). A BLAST search for these subunits in the *H. sulfonivorans* genome has only identified two poor homologues of DdhA and no homologues of DdhB or DdhC. This suggests that if this enzyme does exist in *H. sulfonivorans* then it is distinct from DdhABC.

Although a functional DMSO dehydrogenase has yet to be characterised, the chemical oxidation of DMSO to $DMSO_2$ can occur spontaneously in the presence of a hydroxyl radical

(Miller *et al*., 1996) so it is possible that this reaction may mediate DMSO oxidation *in-vivo* as has been proposed for the oxidation of methanesulfinate (MSIA) to methanesulfonic acid (MSA) (Wicht, 2016).

Dimethylsulfide oxidation pathway

In the DMS oxidation pathway as proposed by Boden *et al.* (2011), DMS is oxidised by a DMS monooxygenase to generate MT and formaldehyde, which is further oxidised by MT to generate hydrogen sulfide, hydrogen peroxide and another molecule of formaldehyde.

An $FMNH_2$-dependent DMS monooxygenase has already been characterised in *Hyphomicrobium sulfonivorans* S1 by Boden *et al.* (2011), consisting of the monooxygenase subunit (DmoA) and FMN oxidoreductase subunit (DmoB). The DmoA encoding gene with NCBI accession number ADU77278.1, has been identified in the new genome by BLAST search as C6Y62_13210. This gene, annotated as a 5,10-methylene H_4MPT reductase, has 99% sequence identity to the original *dmoA* sequence. The DmoB subunit of this DMS monooxygenase has yet to be identified and experimentally characterised, but a potential candidate can be found just upstream as the NCBI annotated "flavin reductase" C6Y62_13200, as was also found by Boden *et al.* (2011).

BLAST analysis of DmoA against the *H. sulfonivorans* genome also identified another putative monooxygenase that may be of interest, C6Y62_13615, also annotated as a 5,10-methylene H_4MPT reductase. This weaker homologue only has a 51% sequence identity to DmoA and a functional analysis by the KEGG BlastKOALA service has annotated the gene as long-chain alkane monooxygenase (K20938), distinct from alkanesulfonate monooxygenases (Li *et al.*, 2008), but it may be useful to observe the relationship between this gene and MSC metabolism using comparative omics.

The other enzyme of the pathway found in *Hyphomicrobium* species is the MT oxidase MtoX, often annotated as a member of the selenium binding protein family, that was originally characterised in *Hyphomicrobium* sp. VS (Eyice *et al.*, 2017). This enzyme oxidises MT in the presence of oxygen and water to generate hydrogen sulfide, formaldehyde and hydrogen peroxide. However, no significant MtoX candidate has been identified in the *H. sulfonivorans* genome nor in the older assembly discussed in Section 3.3, suggesting that MT oxidation in *H. sulfonivorans* is performed by another enzyme, an as yet uncharacterised MT oxidase.

An alternative for an MT oxidase is an MT methyltransferase, such as the MtsAB characterised in the methanogenic Archaea *Methanosarcina barkeri* (Tallant and Krzycki,

1997, Tallant *et al*., 2001). This MT:Coenzyme M (CoM) methyltransferase can transfer the methyl group from either DMS or MT to CoM, generating methyl-CoM and either MT or hydrogen sulfide respectively. Unfortunately, as with the MT oxidase, *H. sulfonivorans* has no significant homologues for the enzyme's MstA subunit, and only a weak homologue for MtsB subunit that appears to be either a methionine synthase based on NCBI annotation, or a 5-methyl H_4F-homocysteine methyltransferase based on KEGG annotation (K00548).

Dimethylsulfone oxidation pathway

The sequential oxidation of $DMSO_2$ oxidation is an alternate mechanism of dimethyl sulfur compound oxidation in which $DMSO_2$ is oxidised to MSIA by a $DMSO_2$ monooxygenase, resulting in the production of formaldehyde and MSA which can be further oxidised by an MSA monooxygenase to sulfite and another molecule of formaldehyde (Schäfer *et al.*, 2009).

A $DMSO_2$ monooxygenase that has already been characterised in *Pseudomonas fluorescens* is the heterodimeric enzyme SfnFG (Wicht, 2016), consisting of the $FMNH_2$-dependent $DMSO_2$ monooxygenase SfnG and an NADH dependent FMN reductase SfnF subunits. This is likely to generate MSA and formaldehyde. A BLAST search for this enzyme in the *H. sulfonivorans* genome highlights two strong candidates for *sfnG* and two more for *sfnF*, which can be paired together by their respective locations in the *H. sulfonivorans* genome. The first pair is the *sfnG* homologue C6Y62_00885 (*sfnG1*), with 71% homology to the reference sequence from *P. fluorescens* annotated by NCBI as a $DMSO_2$ monooxygenase, and the *sfnF* homologue C6Y62_00880 (*sfnF1*), with 45% homology and labelled as encoding an MsuE-type FMN reductase. The second pair, also labelled as a $DMSO_2$ monooxygenase and FMN reductase respectively, are the *sfnG* homologue C6Y62_13190 (*sfnG2*) and *sfnF* homologue C6Y62_13185 (*sfnF2*). Due to the strong homology of these enzymes to the SfnFG from *P. fluorescens* and successful KEGG for each SfnF (K00299) and SfnG (K17228) candidate, both pairs will be tentatively assigned the function of a $DMSO_2$ monooxygenase for the purpose of comparative omics analysis.

The next enzyme in this potential pathway is an MSA monooxygenase, which oxidises the MSA to yield formaldehyde and sulfite. Two bacterial enzymes that catalyse this reaction are the MsmABCD characterised in *Methylosulfonomonas methylovora* (De Marco *et al*., 1999) and the MsuDE characterised in *Pseudomonas aeruginosa* (Kertesz *et al*., 1999). Although a BLAST search of the *H. sulfonivorans* shows no significant homologues of any of

the components of the MsmABCD system, suggesting that the complex is absent in this organism, several homologues were identified for members of the MsuDE MSA monooxygenase. A homologue for the MsuD subunit in *H. sulfonivorans* with 67% sequence identity is C6Y62_00835, labelled as an alkanesulfonate monooxygenase by both NCBI annotation and KEGG BlastKOALA (K04091). Although the C6Y62_00880 and C6Y62_13185 genes have been identified as a potential homologue of the MsuE subunit in the *H. sulfonivorans* genome, both have already been assigned as candidates for the two SfnF FMN reductases described above. However, due to the high sequence similarity between the various monooxygenases and FMN reductases described above it may simply be the case that FMN reductases are promiscuous, each capable of donating $FMNH_2$ to various different monooxygenases. Furthermore, a BLAST search of the alkanesulfonate monooxygenase SsuDE from *E. coli* (Eichhorn *et al*., 1999) produced the same top BLAST hits in *H. sulfonivorans* as the aforementioned MsuDE from *Pseudomonas*, albeit with a slightly lower score, identity and E-value. On this basis C6Y62_00835 will be used as a putative MsuD and the other SfnFG homologues will remain so for further omics analysis unless conflicting with new data.

DMS production pathway

Although the focus of MSC analysis in *H. sulfonivorans* is on MSC degradation and assimilation rather than production, the bacterial production of DMS has been found to be widespread in soil environments (Carrión *et al*., 2015), notably in *Pseudomonas* species. Although DMS production has not been recorded in *Hyphomicrobium* species, it may still be important to confirm the presence or absence of this system in *H. sulfonivorans* for the purpose of building a metabolic model of MSC metabolism.

The methionine-dependent DMS production pathway in question begins with an L-methionine gamma-lyase, characterised as MdeA in *Pseudomonas putida* (Inoue *et al.*, 1997), that cleaves a methanethiol group from L-methionine to generate ammonia, 2-oxobutanoate and MT. This MT is then methylated to DMS by an MddA-type MT S-methyltransferase, characterised in *Pseudomonas deceptionensis*, that transfers a methyl group from S-adenosyl-L-methionine onto MT to generate DMS and S-adenosyl-L-homocysteine (Carrion *et al*., 2017).

A search for a homologue of MddA in the *H. sulfonivorans* genome returned no significant hits, and although a search for MdeA did return two homologues with ~40%

sequence identity, KEGG annotation by BlastKOALA suggests that they are instead the *metY* homologue C6Y62_09455 (K01740) and *metZ* homologue C6Y62_13360 (K10764). This appears to confirm the absence of the methionine-dependent DMS production system in *H. sulfonivorans*.

Table 3.4-2a: *BLAST search of the* H. sulfonivorans *strain S1 2018 draft for genes of methylated sulfur compound metabolism. Proteins highlighted in* ***bold*** *have been selected as strong homologues.*

METHYLATED SULFUR COMPOUND METABOLISM									
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Interconversion of dimethyl sulfur compounds									
DMSO reductase	subunit A	DmsA	UNIPROT/P18775	*Escherichia coli*	C6Y62_10875	molybdopterin oxidoreductase	28%	225	1e-63
	subunit B	DmsB	UNIPROT/P18776	*Escherichia coli*	C6Y62_03170	ferredoxin family protein	29%	32.3	0.006
	subunit C	DmsC	UNIPROT/P18777	*Escherichia coli*	*no significant hits*				
DMS dehydrogenase	subunit A	DdhA	UNIPROT/Q8GPG4	*Rhodovulum sulfidophilum*	C6Y62_10875	molybdopterin oxidoreductase	25%	78.9	7e-16
					C6Y62_01160	formate dehydrogenase subunit alpha	23%	39.3	0.001
	subunit B	DdhB	UNIPROT/Q8GPG3	*Rhodovulum sulfidophilum*	*no significant hits*				
	subunit C	DdhC	UNIPROT/Q8GPG1	*Rhodovulum sulfidophilum*	*no significant hits*				
Dimethylsulfide oxidation pathway									
DMS monooxygenase large subunit		DmoA	NCBI/ADU77278.1	*Hyphomicrobium sulfonvivorans*	**C6Y62_13210**	**5,10-methylene H_4MPT reductase**	**99%**	**970**	**0.0**
					C6Y62_13615	5,10-methylene H_4MPT reductase	51%	467	1e-161
MT oxidase		MtoX	NCBI/ATJ26742.1	*Hyphomicrobium* sp. VS	*no significant hits*				
Dimethylsulfone oxidation pathway									
MSA monooxygenase	subunit A	MsmA	UNIPROT/Q9X404	*Methylosulfonomonas methylovora*	*no significant hits*				
	subunit B	MsmB	UNIPROT/Q9X405	*Methylosulfonomonas methylovora*	*no significant hits*				
	subunit C	MsmC	UNIPROT/Q9X405	*Methylosulfonomonas methylovora*	*no significant hits*				
	subunit D	MsmD	UNIPROT/Q9X406	*Methylosulfonomonas methylovora*	*no significant hits*				
(continued)									

Table 3.4-2b: *BLAST search of the* H. sulfonivorans *strain S1 2018 draft for genes of methylated sulfur compound metabolism (continued). Proteins highlighted in* ***bold*** *have been selected as strong homologues.*

METHYLATED SULFUR COMPOUND METABOLISM *(continued)*									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Dimethylsulfone oxidation pathway (continued)									
$DMSO_2$ monooxygenase	large subunit	SfnG	UNIPROT/Q3KC85	*Pseudomonas fluorescens*	**C6Y62_00885**	**DMSO monooxygenase SfnG**	**73%**	**554**	**0.0**
					C6Y62_13190	**DMSO monooxygenase SfnG**	**59%**	**439**	**5e-154**
					C6Y62_00835	alkanesulfonate monooxygenase SsuD	30%	132	1e-35
	small subunit	SfnF	UNIPROT/Q3K9A2	*Pseudomonas fluorescens*	**C6Y62_00880**	**FMN reductase MsuE**	**45%**	**131**	**1e-38**
					C6Y62_13185	**FMN reductase**	**41%**	**115**	**1e-32**
MSA monooxygenase	large subunit	MsuD	UNIPROT/Q9I1C2	*Pseudomonas aeruginosa*	**C6Y62_00835**	**alkanesulfonate monooxygenase SsuD**	**67%**	**488**	**1e-172**
					C6Y62_00885	DMSO monooxygenase SfnG	32%	119	7e-31
					C6Y62_13190	DMSO monooxygenase SfnG	33%	112	1e-28
	small subunit	MsuE	UNIPROT/Q88J85	*Pseudomonas putida*	C6Y62_00880	FMN reductase MsuE	47%	139	5e-42
					C6Y62_13185	FMN reductase	40%	132	2e-39
alkanesulfonate monooxygenase	large subunit	SsuD	UNIPROT/Q8FJ93	*Escherichia coli*	C6Y62_00835	alkanesulfonate monooxygenase SsuD	62%	479	8e-169
					C6Y62_13190	DMSO monooxygenase SfnG	30%	127	1e-33
					C6Y62_00885	DMSO monooxygenase SfnG	31%	107	4e-27
	small subunit	SsuE	UNIPROT/P80644	*Escherichia coli*	C6Y62_00880	FMN reductase MsuE	37%	76.6	3e-18
					C6Y62_13185	FMN reductase	30%	58.5	4e-12
DMS production pathway									
MT S-methyltransferase		MddA	UNIPROT/ A0A0F6P9C0	*Pseudomonas deceptionensis*	*no significant hits*				
MT:CoM methyltransferase	subunit A	MtsA	UNIPROT/Q48924	*Methanosarcina barkeri*	*no significant hits*				
	subunit B	MtsB	UNIPROT/Q48925	*Methanosarcina barkeri*	C6Y62_09455	methionine synthase	31%	84.3	4e-19
L-methionine gamma-lyase		MdeA	UNIPROT/P13254	*Pseudomonas_putida*	C6Y62_14670	O-acetylhomoserine aminocarboxypropyltransferase	40%	277	2e-89
					C6Y62_05070	O-succinylhomoserine sulfhydrylase MetZ	39%	241	1e-75
					C6Y62_13360	cystathionine beta-lyase MetC	31%	40.8	1e-04
					C6Y62_15465	5-aminolevulinate synthase HemA	33%	38.9	5e-04

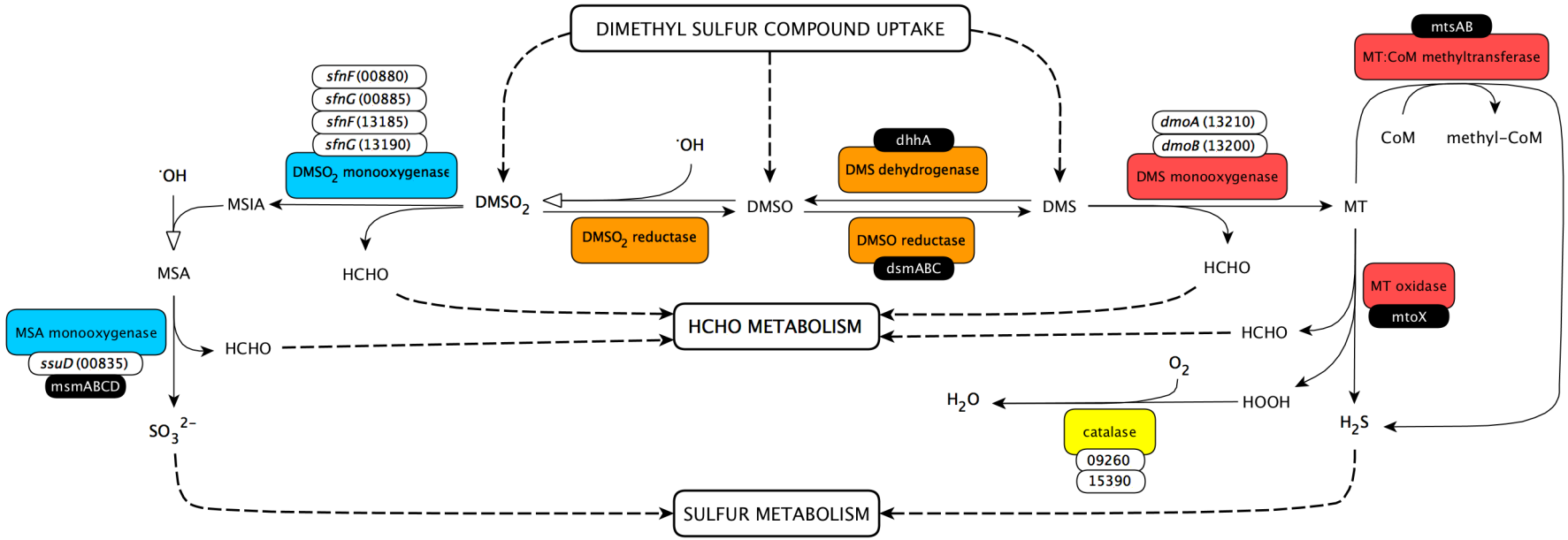


Figure 3.4-2: *Comparative genomics analysis of MSC metabolism in* H. sulfonivorans. General legend for metabolic pathways: Enzymes with homologues in the *H. sulfonivorans* genome have been annotated with a WHITE label, containing the gene name and ID. Genes with low confidence are marked with a (*) and dotted outline. Genes with no homologue in the *H. sulfonivorans* genome are labelled BLACK. Mechanisms of MSC metabolism are labelled as follows: Interconversion of DMSCs (ORANGE), DMS oxidation pathway (RED), $DMSO_2$ oxidation pathway (BLUE), hydrogen peroxide detoxification (YELLOW). Non-enzymatic reactions are marked by a WHITE arrowhead.

3.4.3 Formaldehyde Metabolism

Formaldehyde is a highly reactive, toxic intermediate of methylotrophic C1 compound metabolism that can either be dissimilated to CO_2 or assimilatied by a variety of mechanisms depending on the particular type of methylotroph involved. As *H. sulfonivorans* is a serine cycle methylotroph (Boden *et al.*, 2011), we would expect formaldehyde derived from MSC or MeOH metabolism to be used to generate 5,10-methylene H_4F, the methylated H_4F cofactor used as the entry point for carbon assimilation via the serine cycle (Chistoserdova, 2011). Although 5,10-methylene H_4F can be formed by the spontaneous reaction of formaldehyde with H_4F, methylotrophic bacterial species may also generate the compound enzymatically via the C1 intermediate formate (HCOOH) (Chistoserdova *et al.*, 2003).

In either assimilatory or dissimilatory formaldehyde metabolism, the first step is to is convert formaldehyde to formate by either the tetrahydromethanopterin (H_4MPT)-dependent formaldehyde activation pathway, glutathione (GSH)-dependent formaldehyde activation pathway, or oxidation via a formaldehyde dehydrogenase (Chistoserdova, 2011). Formate dissimilation can then occur by formate oxidation via a formate dehydrogenase, producing CO_2 for diffusion out of the cell. Carbon assimilation from formate has been shown to occur in certain Serine cycle methylotrophs via the H_4F pathway (Crowther *et al.*, 2008), in which formate is used to methylate tetrahydrofolate (H_4F) to yield 5,10-methylene H_4F. The results of a BLAST search for common enzymes of methylotrophic MSC metabolism in the *H. sulfonivorans* genome are displayed in Tables 3.4-3 (a, b), while the successfully identified enzymes of formaldehyde metabolism have been mapped onto a putative metabolic pathway in Figure 3.4-3.

Formaldehyde oxidation

H. sulfonivorans has two potential homologues for a formaldehyde dehydrogenase in the putative alcohol dehydrogenases C6Y62_10250 and C6Y62_02365, which respectively have 36% and 26% homology to the FdhA of *Pseudomonas putida* (Ito *et al*., 1994) . As the stronger homologue, C6Y62_10250, has also been annotated by BlastKOALA as the 2-propanol oxidising alcohol dehydrogenase, *adh2* (K18369), the function of this enzyme remains uncertain and it is possible that *H. sulfonivorans* has no functional homologue of FdhA. However, for speculative purposes C6Y62_10250 and C6Y62_02365 will be carried forward for further omics analysis as weak candidates for FdhA.

Table 3.4-3a: *BLAST search of the* H. sulfonivorans *strain S1 2018 draft for genes of formaldehyde metabolism. Proteins highlighted in* ***bold*** *have been selected as strong homologues.*

FORMALDEHYDE METABOLISM									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Formaldehyde oxidation									
formaldehyde dehydrogenase		FdhA	UNIPROT/P46154	*Pseudomonas putida*	C6Y62_10250	alcohol dehydrogenase	36%	102	3e-25
					C6Y62_02365	alcohol dehydrogenase	26%	37.0	0.001
Glutathione pathway									
GSH-dependent formaldehyde activating enzyme		Gfa	UNIPROT/Q51669	*Paracoccus denitrificans*	C6Y62_08655	GFA family protein	24%	37.4	2e-04
					C6Y62_13875	GFA family protein	28%	36.6	2e-04
hydroxymethyl-GSH dehydrogenase		FlhA	UNIPROT/P45382	*Paracoccus denitrificans*	C6Y62_10250	alcohol dehydrogenase	26%	61.6	1e-11
Tetrahydromethanopterin pathway									
formaldehyde-activating enzyme		Fae	UNIPROT/Q9FA38	*Methylobacterium extorquens*	**C6Y62_12595**	**formaldehyde-activating enzyme Fae**	**76%**	**275**	**1e-95**
					C6Y62_05220	aldehyde-activating protein	32%	79.7	1e-19
					C6Y62_13025	formaldehyde-activating enzyme Fae	32%	73.9	1e-17
methylene-H_4MPT dehydrogenase A		MtdA	UNIPROT/P55818	*Methylobacterium extorquens*	C6Y62_12620	methylene-H_4MPT dehydrogenase	32%	58.5	8e-11
NADP-dependent methylene-H_4MPT dehydrogenase B		MtdB	UNIPROT/O85012	*Methylobacterium extorquens*	**C6Y62_12620**	**methylene-H_4MPT dehydrogenase**	**53%**	**266**	**3e-88**
methenyl-H_4MPT cyclohydrolase		Mch	UNIPROT/O85014	*Methylobacterium extorquens*	**C6Y62_12610**	**methenyl-H_4MPT cyclohydrolase**	**58%**	**349**	**6e-120**
formyltransferase/ hydrolase complex	subunit A	FhcA	UNIPROT/C5B137	*Methylobacterium extorquens*	**C6Y62_12510**	**formyl-MFR dehydrogenase subunit A**	**40%**	**404**	**1e-134**
	subunit B	FhcB	UNIPROT/C5B138	*Methylobacterium extorquens*	**C6Y62_12515**	**formyl-MFR dehydrogenase**	**32%**	**59.7**	**9e-11**
	subunit C	FhcC	UNIPROT/C5B135	*Methylosulfonomonas methylovora*	**C6Y62_12500**	**Formyl-MFR dehydrogenase subunit C**	**43%**	**162**	**7e-49**
	Subunit D	FhcD	UNIPROT/Q49118	*Methylosulfonomonas methylovora*	**C6Y62_12505**	**Formyl-MFR H_4MPT N-formyltransferase**	**59%**	**346**	**3e-119**
(continued)									

Table 3.4-3b: *BLAST search of the* H. sulfonivorans *strain S1 2018 draft for genes of formaldehyde metabolism (continued). Proteins highlighted in* ***bold*** *have been selected as strong homologues.*

FORMALDEHYDE METABOLISM *(continued)*								
REFERENCE SEQUENCE				BLAST RESULTS				
Enzyme	Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Formate oxidation								
formate dehydrogenase	Fdh	UNIPROT/P33160	*Pseudomonas* sp. 101	**C6Y62_05595**	**NAD-dependent formate dehydrogenase**	**80%**	**672**	**0.0**
				C6Y62_02980	**NAD-dependent formate dehydrogenase**	**75%**	**623**	**0.0**
				C6Y62_01355	D-glycerate dehydrogenase	29%	116	4e-30
				C6Y62_15775	D-glycerate dehydrogenase	31%	113	4e-29
Tetrahydrofolate pathway								
formate-THF ligase	Fhs	UNIPROT/Q83WS0	*Methylobacterium extorquens*	**C6Y62_04850**	**formate H_4F ligase**	**65%**	**675**	**0.0**
bifunctional methenyl-H_4F cyclohydrolase/ methylene-H_4F dehydrogenase	FolD	UNIPROT/P24186	*Escherichia coli*	**C6Y62_04710**	**bifunctional methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase FolD**	**55%**	**280**	**8e-94**

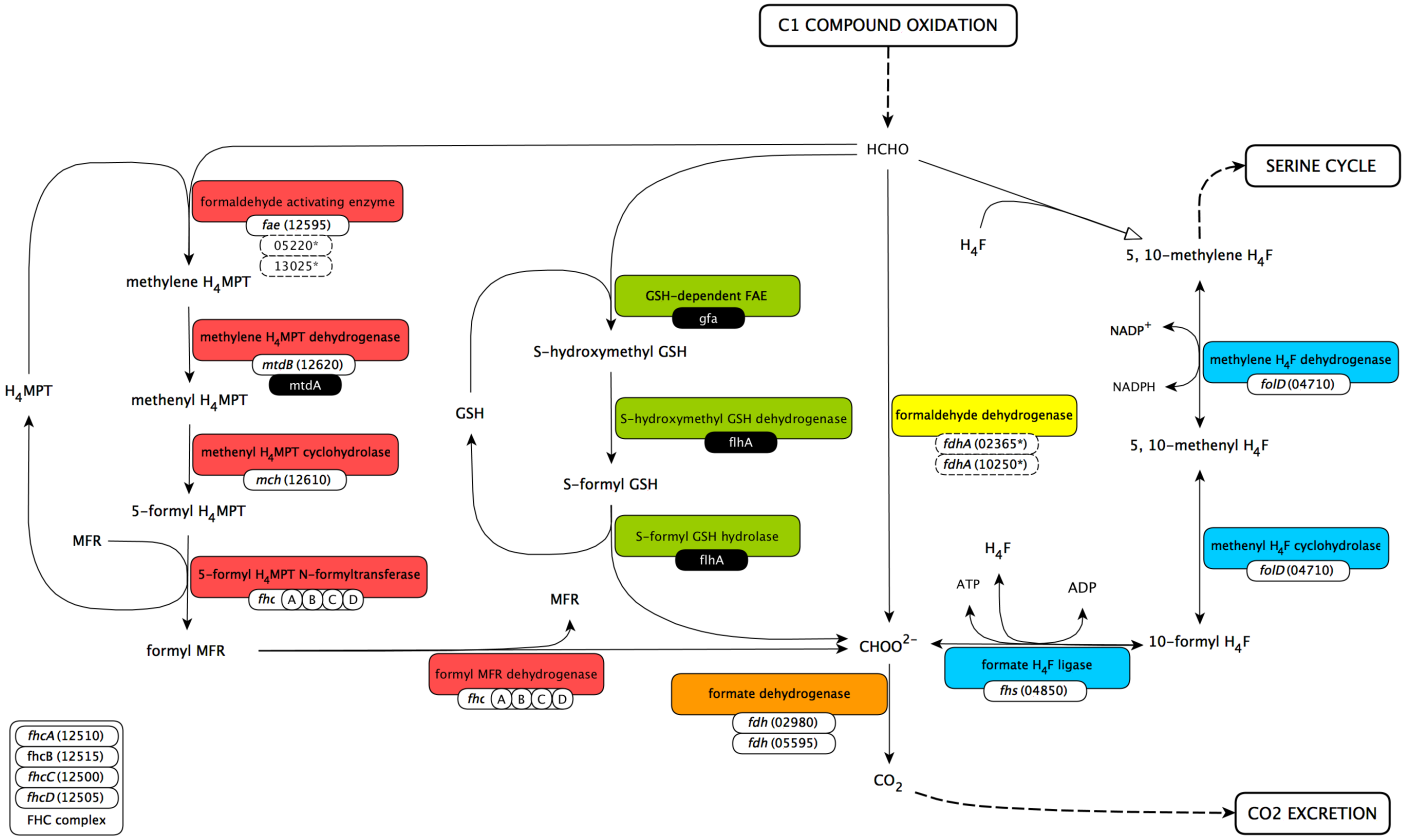


Figure 3.4-3: *Comparative genomics analysis of formaldehyde metabolism in* H. sulfonivorans. General legend for metabolic pathways: Enzymes with homologues in the *H. sulfonivorans* genome have been annotated with a WHITE label, containing the gene name and ID. Genes with low confidence are marked with a (*) and dotted outline. Genes with no homologue in the *H. sulfonivorans* genome are labelled BLACK. Mechanisms of formaldehyde (HCHO) metabolism are labelled as follows: H_4MPT-dependent pathway (RED), GSH-dependent pathway (GREEN), formaldehyde oxidation (YELLOW), formate oxidation (ORANGE) and H_4F pathway (BLUE). Non-enzymatic reactions are marked by a WHITE arrowhead.

Glutathione pathway

The GSH-dependent formaldehyde activation pathway is a mechanism of producing formate from formaldehyde seen in certain species of methylotrophic bacteria, such as the Alphaproteobacteria *Paracoccus denitrificans* (Goenrich *et al.,* 2002)*,* that has been used as a source of reference sequences for BLAST searches of the *H. sulfonivorans* genome

The first enzyme is a GSH-dependent formaldehyde activating enzyme, GFA, which activated formaldehyde in the presence of GSH to produce S-hydroxymethyl GSH. Two poor homologues for this enzyme have been found in the *H. sulfonivorans* genome, C6Y62_08655 and C6Y62_13875, showing 24% and 28% sequence identity respectively and both of which have been given the NCBI annotation of a GFA-like enzyme.

The second enzyme is a bifunctional S-hydroxymethyl GSH dehydrogenase/ S-formyl GSH hydrolase, FlhA, that oxidises S-hydroxymethyl GSH to S-formyl GSH, then cleaves S-formyl GSH to yield formate and regenerate GSH. This has only one weak homologue in the genome, C6Y62_10250 with an identity of 26% that is annotated as an alcohol dehydrogenase. As C6Y62_10250 has also been successfully annotated by KEGG BlastKOALA as the 2-propanol oxidising alcohol dehydrogenase, Adh2 (K18369), it appears that FlhA is absent from the *H. sulfonivorans* genome.

Tetrahydromethanopterin pathway

The H_4MPT-dependent formaldehyde activation pathway that converts formaldehyde to formate consists of five reactions. BLAST searches for the enzymes catalysing these reactions in the *H. sulfonivorans* genome have been taken from the model Alphaproteobacteria *Methylobacterium extorquens* (Marx *et al*., 2003). Two of these reactions are mediated by a formaldehyde activating enzyme, Fae, activating formaldehyde with H_4MPT to generate methylene H_4MPT. The gene C6Y62_12595 has been identified as a strong candidate for this enzyme as it shares 76% sequence identity, supported by both NCBI and KEGG (K10713) annotation as a Fae-type formaldehyde activating enzyme. Two weaker homologues C6Y62_05220 and C6Y62_13025 have also been identified with only 32% sequence, neither of which have been assigned a KEGG KO identity by BlastKOALA, but will both be taken forward as weak candidates for Fae.

In the next two steps of this pathway methylene-H_4MPT is oxidised to methenyl-H_4MPT by either an methylene-H_4MPT dehydrogenase type A (MtdA) or type B (MtdB), which is then converted to 5-formyl H_4MPT by a methenyl H_4MPT cyclohydrolase (Mch)

(Marx *et al*., 2003, Chistoserdova, 2011). BLASTs of MtdA and MtdB against the *H. sulfonivorans* genome highlighted C6Y62_12620 as the only potential candidate for either an MtdA (32% identity) or MtdB (53% identity), the latter being the more likely two enzymes due based on BLAST score and sequence identity. On this basis *H. sulfonivorans* appears to have an MtdB C6Y62_12620, supported by KEGG annotation (K10714), but lacks an MtdA. An Mch type methenyl H_4MPT cyclohydrolase has also been clearly identified as C6Y62_12610 in the *H. sulfonivorans* genome, supported by a relatively strong sequence identity of 58% to the Mch of *Methylobacterium extorquens* and successful KEGG annotation (K01499).

The final two reactions of the H_4MPT-dependent formaldehyde activation pathway are mediated by the multifunctional Fhc complex, consisting of the subunits FhcA, FhcB, FhcC and FhcD. This complex acts as a 5-formyl H_4MPT N-formyltransferase to transfer the formyl group from 5-formyl H_4MPT to MFR, producing 5-formyl MFR and regenerating H_4MPT, then oxidising formyl MFR to produce formate and regenerate MFR. A single homologue has been identified in *H. sulfonivorans* for each of the four Fhc subunits, FhcA (C6Y62_12510), FhcB (C6Y62_12515), FhcC (C6Y62_12500) and FhcD (C6Y62_12505), which is supported by their four respective KEGG annotations of K00200, K00201, K00202 and K00672.

Formate oxidation

Formate generated by formaldehyde oxidation and H_4MPT-dependent formaldehyde activation can either be used as a substrate for H_4F methylation and assimilation via the serine cycle, or dissimilated by oxidation to CO_2 by a formate dehydrogenase, Fdh. Two strong candidates for this enzyme have been found in the genes C6Y62_05595 and C6Y62_02980, both of which have been annotated as NAD-dependent formate dehydrogenases by NCBI and KEGG (K00122) and have quite close homology to a Fdh examined in *Pseudomonas* species (Tishkov *et al.*, 1993).

Tetrahydrofolate pathway

Methylotrophic carbon assimilation in serine cycle methylotrophs is dependent on the production of the methylated cofactor 5,10-methylene H_4F. Although this compound can be generated from formaldehyde and H_4F in a spontaneous condensation reaction (Escalante-Semerena *et al*., 1984; Vorholt *et al*., 2000), in certain Alphaproteobacteria such as *Methylovorum extorquens* it can also be generated from formate via the H_4F pathway (Crowther *et al*., 2008). Here, an Fhs-type ATP-dependent formate H_4F ligase synthesises

10-formyl H_4F from formate and H_4F. The 10-formyl H_4F can then be converted to 5,10-methenyl H_4F and then oxidised to 5,10-methylene H_4F in two reactions catalysed by the bifunctional methenyl-H_4F cyclohydrolase/ methylene-H_4F dehydrogenase FolD (Marx *et al.*, 2003, Chistoserdova, 2011).

A strong homologue has been identified in *H. sulfonivorans* with 65% sequence identity to the Fhs protein from *M. extorquens*, the NCBI and KEGG (K01938) annotated formate H_4F ligase encoding gene C6Y62_04850. A putative bifunctional methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase has also been identified by KEGG annotation (K01491) and BLAST as homologue C6Y62_04710, showing 55% sequence identity to FolD from *Escherichia coli*.

3.4.4 Inorganic Sulfur Metabolism

Based on the comparative genomics outlined in Section 3.4.3 and previous research performed by Borodina *et al.* (2000, 2002) and Boden *et al*. (2011), we would expect to see the metabolism of DMSCs yield the inorganic sulfur compounds hydrogen sulfide and/or sulfite. As *H. sulfonivorans* can utilise MSCs as both a sole carbon and sole sulfur source the organism is likely to participate in both assimilatory and dissimilatory inorganic sulfur compound metabolism depending on its substrate needs. Bacterial sulfur assimilation is likely to occur via amino acid synthesis; cysteine synthesis is the typical mechanism of assimilatory sulfur metabolism (Kredich, 2008) but homocysteine synthesis has also been observed in other certain Alphaproteobacteria such as *Pseudomonas* species (Vermeij and Kertesz, 1999a). Based on a recent publication by Koch and Dahl (2018) into the dissimilatory DMS metabolism of another *Hyphomicrobium* species, the inorganic sulfur compounds produced from MSCs are likely to be dissimilated via the production and oxidation of thiosulfate, generating sulfate which is excreted from the cell. The results of a BLAST search for common enzymes of methylotrophic MSC metabolism in the *H. sulfonivorans* genome are displayed in Table 3.4-4, while the successfully identified enzymes of inorganic sulfur metabolism have been mapped onto a putative metabolic pathway in Figure 3.4-4.

Cysteine synthesis

Cysteine synthesis typically begins with the acetylation of L-serine to O_3-acetyl-L-serine by an acetyl CoA-dependent serine O-acetyltransferase, which a cysteine synthase then uses as a substrate for the synthesis of L-cysteine from hydrogen sulfide (Sekowska, 2000;

Kredich, 2008). The CoA and acetate respectively released from the first and second steps can then be recycled to produce acetyl CoA by a third enzyme, an acetyl-CoA synthetase.

BLAST searches using the cysteine synthesis system of *E. coli* shows good candidates for two CysK-like cysteine synthases, C6Y62_15815 and C6Y62_08560, which have 53% and 38% sequence identity respectively and are both supported by functional KEGG annotation (K01738). Putative enzymes have also been found for the serine O-acetyltransferase (C6Y62_09280) and acetyl-CoA synthetase (C6Y62_04625) based on NCBI and KEGG annotations (K00640 and K01895 respectively).

In addition to A-type cysteine synthases such as CysE, B-type cysteine synthases (CysM) are able to mediate another O_3-acetyl-L-serine-dependent reaction that leads to cysteine synthesis, in which thiosulfate is used to generate S-sulfocysteine from O_3-acetyl-L-serine (Claus *et al*., 2005). This can then be desulfonated in the presence of either a glutaredoxin (GRX) or thioredoxin (TRX) to generate cysteine, sulfate and either GRX-S_2 or TRX-S_2. Although the two putative cysteine synthases C6Y62_15815 and C6Y62_08560 appear to be homologues of the cysteine synthase B (CysM) from *E. coli*, they both appear to be more closely related to CysE and it is most likely that they are both A-type cysteine synthases.

Table 3.4-4a: *BLAST search of the* H. sulfonivorans *strain S1 2018 draft for genes of inorganic sulfur compound metabolism.*

INORGANIC SULFUR METABOLISM								
REFERENCE SEQUENCE				BLAST RESULTS				
Enzyme	Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Sulfite metabolism								
sulfite reductase	CysJ	UNIPROT/P38038	*Escherichia coli*	**C6Y62_12675**	**sulfite reductase subunit alpha**	**38%**	**365**	**2e-118**
				C6Y62_15160	flavodoxin/nitric oxide synthase	26%	38.5	3e-04
SOE, molybdopterin oxidoreductase alpha subunit	SoeA	UNIPROT/D3RNN8	*Allochromatium vinosum*	C6Y62_10875	molybdopterin oxidoreductase	26%	88.2	9e-19
				C6Y62_05050	CbbBc protein	26%	53.9	4e-08
SOE, ferredoxin iron-sulfur binding domain protein	SoeB	UNIPROT/D3RNN7	*Allochromatium vinosum*	C6Y62_00195	cytochrome c oxidase accessory protein CcoG	29%	33.9	0.007
SOE, DMSO reductase anchor subunit	SoeC	UNIPROT/D3RNN6	*Allochromatium vinosum*	*no significant hits*				
Cysteine synthesis								
serine O-acetyltransferase	CysK/ MetA	UNIPROT/ A0A1D3PCK2	*Lactobacillus acidophilus*	C6Y62_15815	cysteine synthase A, CysK	38%	160	7e-47
serine acetyltransferase	CysE	UNIPROT/P0A9D4	*Escherichia coli*	C6Y62_09280	serine O-acetyltransferase	56%	281	7e-95
				C6Y62_14330	N-acetyltransferase	31%	44.7	1e-06
cysteine synthase A	CysK	UNIPROT/P0ABK5	*Escherichia coli*	**C6Y62_15815**	**cysteine synthase A, CysK**	**53%**	**251**	**1e-81**
				C6Y62_08560	**cysteine synthase A**	**38%**	**145**	**7e-41**
cysteine synthase B	CysM	UNIPROT/P16703	*Escherichia coli*	C6Y62_15815	cysteine synthase A, CysK	38%	160	7e-47
				C6Y62_08560	cysteine synthase A	34%	130	8e-36
				C6Y62_06910	threonine ammonia-lyase	24%	36.2	0.002
				C6Y62_05590	HPP family protein	30%	34.3	0.004
Homocysteine synthesis								
O-succinylhomoserine sulfhydrylase	MetZ	UNIPROT/P55218	*Pseudomonas aeruginosa*	**C6Y62_05070**	**O-succinylhomoserine sulfhydrylase**	**50%**	**341**	**3e-114**
				C6Y62_14670	O-acetylhomoserine aminocarboxypropyltransferase	38%	253	5e-80
				C6Y62_13360	cystathionine beta-lyase	31%	140	5e-38
Thiosulfate dehydrogenation								
thiosulfate dehydrogenase	TsdA	UNIPROT/D3RVD4	*Allochromatium vinosum*	*no significant hits*				
(continued)								

Table 3.4-4b: *BLAST search of the* H. sulfonivorans *strain S1 2018 draft for genes of inorganic sulfur compound metabolism (continued). Proteins highlighted in* ***bold*** *have been selected as strong homologues.*

INORGANIC SULFUR METABOLISM *(continued)*									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Sox-mediated thiosulfate oxidation									
SOX complex	subunit A	SoxA	UNIPROT/O33434	*Paracoccus pantotrophus*	*no significant hits*				
	subunit X	SoxX	UNIPROT/ A0A1I5IPQ3	*Paracoccus pantotrophus*	*no significant hits*				
	subunit B	SoxB	UNIPROT/ A0A1K2FBC2	*Paracoccus pantotrophus*	*no significant hits*				
	subunit C	SoxC	UNIPROT/A0A1I5IPK4	*Paracoccus pantotrophus*	**C6Y62_13220**	**sulfite dehydrogenase**	**42%**	**308**	**8e-101**
					C6Y62_09385	molybdopterin-binding protein	26%	53.9	4e-09
	subunit D	SoxD	UNIPROT/A0A1I5INZ2	*Paracoccus pantotrophus*	**C6Y62_13225**	**cytochrome C**	**43%**	**133**	**2e-37**
					C6Y62_15940	cytochrome c family protein	30%	93.2	8e-24
					C6Y62_15140	cytochrome C	46%	89.4	6e-22
					C6Y62_12640	cytochrome c family protein	52%	81.3	2e-19
	subunit Y	SoxY	UNIPROT/A0A1I5KIK2	*Paracoccus pantotrophus*	C6Y62_14215	quinoprotein dehydrogenase-associated SoxYZ-like carrier	43%	162	1e-48
					C6Y62_12280	quinoprotein dehydrogenase-associated SoxYZ-like carrier	34%	130	7e-37
	subunit Z	SoxZ	UNIPROT/ A0A089NU96	*Methylobacterium oryzae*	C6Y62_14215	quinoprotein dehydrogenase-associated SoxYZ-like carrier	46%	192	4e-60
					C6Y62_12280	quinoprotein dehydrogenase-associated SoxYZ-like carrier	34%	147	4e-43
	subunit Y	SoxY	UNIPROT/D8JT39	*Hyphomicrobium denitrificans*	**C6Y62_10660**	**thiosulfate oxidation carrier protein SoxY**	**50%**	**105**	**8e-30**
	subunit Z	SoxZ	UNIPROT/D8JX75	*Hyphomicrobium denitrificans*	**C6Y62_10665**	**thiosulfate oxidation carrier complex protein SoxZ**	**42%**	**95.5**	**7e-27**
(continued)									

Table 3.4-4c: *BLAST search of the* H. sulfonivorans *strain S1 2018 draft for genes of inorganic sulfur compound metabolism (continued). Proteins highlighted in* ***bold*** *have been selected as strong homologues.*

INORGANIC SULFUR METABOLISM *(continued)*									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Hdr-mediated thiosulfate oxidation									
heterodisulfide reductase	subunit A	HdrA	UNIPROT/D8JT26	*Hyphomicrobium denitrificans*	*no significant hits*				
	subunit B	HdrB	UNIPROT/ A0A088STN3	*Acidithiobacillus caldus*	C6Y62_13690	phosphoribosylpyrophosphate synthetase	24%	33.9	0.007
	subunit C	HdrC	UNIPROT/ A0A088S946	*Acidithiobacillus caldus*	C6Y62_11890	glycolate oxidase iron-sulfur subunit	29%	43.5	5e-06
					C6Y62_08700	NADH-quinone oxidoreductase subunit	28%	35.8	9e-04
					C6Y62_13005	Fe-S oxidoreductase	23%	30.4	0.005
sulfur carrier protein		TusA	UNIPROT/P0A890	*Escherichia coli*	*no significant hits*				
lipoate binding protein		LbpA	UNIPROT/D8JT31	*H. denitrificans*	*no significant hits*				
Sulfate adenylation pathway									
adenylyl-sulfate kinase		CysC	UNIPROT/P57702	*Pseudomonas aeruginosa*	**C6Y62_15095**	**sulfate adenylyltransferase subunit CysC**	**51%**	**167**	**6e-50**
sulfate adenylyl-transferase	subunit 1	CysN	UNIPROT/P23845	*Escherichia coli*	C6Y62_12695	sulfate adenylyltransferase	43%	333	4e-109
	subunit 2	CysD	UNIPROT/O50273	*Pseudomonas aeruginosa*	**C6Y62_12700**	**sulfate adenylyltransferase subunit CysD**	**62%**	**410**	**2e-144**
bifunctional sulfate adenylyl-transferase/ adenylyl-sulfate kinase	subunit 1	CysCN	UNIPROT/O50274	*Pseudomonas aeruginosa*	**C6Y62_12695**	**sulfate adenylyltransferase**	**42%**	**333**	**2e-107**
					C6Y62_15095	adenylyl-sulfate kinase, cysC	35%	155	5e-42
					C6Y62_11295	elongation factor Tu	29%	90.1	3e-20
					C6Y62_12105	peptide chain release factor 3	35%	52.8	4e-08
phosphoadenosine phosphosulfate reductase		CysH	UNIPROT/O05927	*Pseudomonas aeruginosa*	**C6Y62_12705**	**phosphoadenylyl-sulfate reductase**	**37%**	**157**	**5e-47**

Table 3.4-4d: *BLAST search of the* H. sulfonivorans *strain S1 2018 draft for genes of inorganic sulfur compound metabolism (continued). Proteins highlighted in* ***bold*** *have been selected as strong homologues.*

INORGANIC SULFUR METABOLISM *(continued)*									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Sulfide oxidation									
sulfide:quinone oxidoreductase		SqrB	UNIPROT/D8JTH6	*Hyphomicrobium denitrificans*	**C6Y62_00310**	**TIGR01244 family phosphatase**	**66%**	**544**	**0.0**
Persulfide dioxygenase		Pdo	UNIPROT/D8JTH6	*Hyphomicrobium denitrificans*	**C6Y62_00305**	**MBL fold metallo-hydrolase**	**66%**	**491**	**9e-177**
rhodanese-type sulfur transferase		Rhd	UNIPROT/D8JTH7	*Hyphomicrobium denitrificans*	**C6Y62_00310**	**TIGR01244 family phosphatase**	**52%**	**120**	**2e-33**
sulfide dehydrogenase	subunit A	FccA	UNIPROT/Q06529	*Allochromatium vinosum*	**C6Y62_05645**	**cbb3-type cytochrome-c oxidase**	**67%**	**34.3**	**0.004**
	subunit B	FccB	UNIPROT/Q06530	*Allochromatium vinosum*	C6Y62_10655	flavocytochrome C	44%	263	2e-83

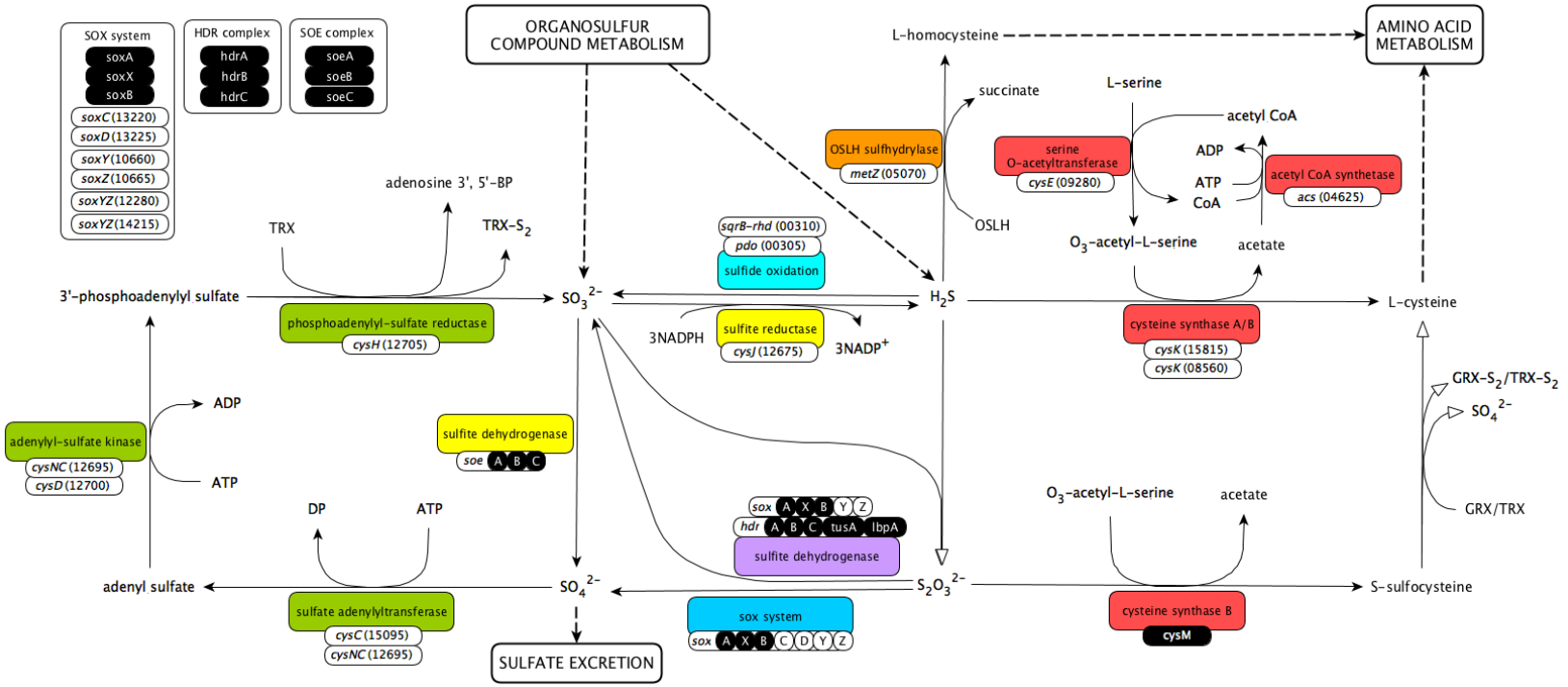


Figure 3.4-4: *Comparative genomics analysis of inorganic sulfur compound MSC metabolism in* H. sulfonivorans. Note the uncommon abbreviations OSLH (O-succinyl-L-homoserine), TRX (thioredoxin), GRX (glutaredoxin) and DP (diphosphate). General legend for metabolic pathways: Enzymes with homologues in the *H. sulfonivorans* genome have been annotated with a WHITE label, containing the gene name and ID. Genes with no homologue in the *H. sulfonivorans* genome are labelled BLACK. Mechanisms of inorganic sulfur compound metabolism are labelled as follows: sulfite metabolism (YELLOW), cysteine synthesis (RED), homocysteine synthesis (ORANGE), Sox-mediated thiosulfate oxidation (BLUE), Hdr-mediated thiosulfate oxidation (PURPLE), sulfate adenylylation pathway (GREEN). Non-enzymatic reactions are marked by a WHITE arrowhead.

Homocysteine synthesis

Another mechanism of bacterial sulfur compound assimilation that has been identified in certain Alphaproteobacteria is homoserine synthesis, which uses hydrogen sulfide and O-succinylhomoserine (OSHS) to generate homocysteine and succinate (Vermeij and Kertesz, 1999). In *Pseudomonas* species, this process is catalysed by the O-succinylhomoserine sulfhydrylase MetZ (Foglino *et al*., 1995). A BLAST search for this enzyme in *H. sulfonivorans* shows that the gene C6Y62_05070 is a likely *metZ* candidate, based on homology with a sequence identity of 50% NCBI and KEGG annotation (K10764), suggesting the possibility that this mechanism of sulfur assimilation also appears in *H. sulfonivorans*.

Sox-mediated thiosulfate oxidation

The multi-enzyme Sox system mediates the bacterial oxidation of thiosulfate to sulfate, and can be split into the four distinct enzymes: SoxAX, SoxB, SoxCD and SoxYZ (Friedrich *et al.*, 2005a), all of which have previously been identified in *Hyphomicrobium denitrificans* by Koch and Dahl (2018). However, it appears that *H. sulfonivorans* may have only a partial Sox system consisting of SoxCD and SoxYZ as outlined below.

A BLAST search for the Sox enzymes from *Paracoccus pantotrophus* in the *H. sulfonivorans* genome has yielded no significant candidates for the SoxA, SoxB or SoxX proteins. As no predicted proteins have been identified in this organism for SoxA (K17222), SoxB (K17223) or SoxX (K17224) by functional KEGG annotation either, this suggests that the SoxAX and SoxB are missing from the organism. However, a search for SoxCD has identified the SoxC homologue C6Y62_13220 and SoxD homologue C6Y62_13225, which have the respective KEGG annotation/ sequence identity of K17225/ 42% (SoxC) and K22622/ 43% (SoxD).

Three potential *soxY* candidates have been found by searching the *H. sulfonivorans* genome based on KEGG annotation (K17226). These are C6Y62_10660, C6Y62_14215 and C6Y62_12280, while a search for *soxZ* (K17227) found the potential candidates C6Y62_10665. Two of these genes, C6Y62_14215 and C6Y62_12280, are annotated as encoding quinoprotein dehydrogenase-associated SoxYZ-like carrier proteins and have been identified as homologues of both SoxY from *Paracoccus pantotrophus* and SoxZ from *Methylobacterium oryzae*, suggesting that they are SoxYZ-like fusion proteins, though regrettably members of this family have yet to be characterised. Further BLAST searches using the putative SoxY and SoxZ from *H. denitrificans* (Koch and Dahl, 2018) identified the third

soxY candidate C6Y62_10660 and the KEGG annotated *soxZ*, C6Y62_10665. Given the NCBI and KEGG annotations of these genes and their close proximity in the genome, they have been taken forward as putative genes encoding SoxY (C6Y62_10660), SoxZ (C6Y62_10665) and SoxYZ (C6Y62_12280, C6Y62_14215).

Hdr-mediated thiosulfate oxidation

The heterodisulfide reductase mediated oxidation of thiosulfate to sulfite is another mechanism of thiosulfate oxidation that has been studied in *Hyphomicrobium denitrificans*, thought to involve a multi-enzyme HdrC1B1AHypHdrC2B2 complex (Koch and Dahl, 2018). This complex, together with the sulfur carrier TusA, lipoate binding protein LbpA, various membrane transporters and members of the Sox system SoxAXBYZ, has been suggested by Koch and Dahl (2018) to oxidise periplasmic thiosulfate to cytoplasmic sulfite.

BLAST searches have been performed for HdrA, HdrB and HdrC in *H. sulfonivorans* using the same sequences as Koch and Dahl (2018) used to identify the Sox system in *H. denitrificans* but have failed to identify any significant candidates for these proteins.

Further BLAST searches for the HdrABC associated proteins TusA from *E. coli* (Dahl *et al*., 2013) and LdrA from *H. denitrificans* (Koch and Dahl, 2018), have yielded only poor homologues with unrelated KEGG annotations. This suggests that the Hdr-mediated thiosulfate oxidation system, and potentially the *hdr* gene cluster as described in *H. denitrificans*, is absent from the *H. sulfonivorans* genome.

Sulfate adenylylation pathway

The sulfate adenylylation pathway mediates the ATP-dependent reduction of sulfate to sulfite, typically associated with assimilatory sulfate metabolism (Sekowska, 2000; Pinto *et al*., 2004). The pathway begins by the adenylation of folate by an ATP-dependent sulfate adenylyltransferase to produce adenylyl-sulfate (APS) and diphosphate (DP), typically mediated by either CysND or a bifunctional CysNCD. The adenylyl-sulfate is then phosphorylated to 3'-phosphoadenylyl-sulfate (PAPS) by an ATP-dependent adenylyl-sulfate kinase, such as CysC or CysNCD, and the 3'-phosphoadenylyl-sulfate reduced to sulfide and adenosine 3',5'-bisphosphate by the thioredoxin-dependent phosphoadenosine phosphosulfate reductase CysH.

A BLAST search for these enzymes using homologues from *Pseudomonas aeruginosa* suggest the presence of *cysC* (C6Y62_15095), *cysD* (C6Y62_12700) and *cysH* (C6Y62_12705), supported by functional KEGG BlastKOALA BLAST for K00955, K00957 and K00390 respectively. A BLAST search for CysN from *Escherichia coli* and CysNC from *Pseudomonas aeruginosa* highlighted the presence of the putative CysNC encoding gene C6Y62_12695. Although this gene returns the KEGG orthology identifier of CysC (K00955) rather than a CysN (K00956), this is not unlikely given that CysNC proteins are a fusion protein of CysC and CysN (Pinto *et al*., 2004). Together, this suggests a functional sulfate adenylylation pathway may be present in *H. sulfonivorans*.

Sulfite metabolism

The typical mechanisms of bacterial sulfite metabolism involve the reduction of sulfite to hydrogen sulfide by a sulfite reductase, sulfide reduction to sulfite by a sulfide dehydrogenase or similar mechanism, sulfite oxidation to sulfate by a sulfite oxidase, and/or the spontaneous reaction of sulfite with hydrogen sulfide, persulfide or sulfur to generate thiosulfate (Heunisch, 1977 via Koch and Dahl, 2018). As described above, the production of hydrogen sulfide is typically associated with assimilatory sulfur metabolism, while the production of sulfite, sulfate and thiosulfate is associated with dissimilatory sulfur metabolism.

Beginning with sulfite reduction, a BLAST and KEGG orthology search has revealed a good candidate for the NADPH-dependent CysJ sulfite reductase (K00380) in the *H. sulfonivorans* genome as C6Y62_12675, which has been labelled by NCBI annotation as the sulfite reductase subunit alpha.

The sulfite oxidase family is a group of molybdoenzymes that oxidise sulfite with oxygen and water, to generate sulfate and hydrogen peroxide (Kappler, 2011). A BLAST search for the subunits of the sulfite family enzyme, SoeABC, has only detected poor homologues of the SoeA and SoeB subunits with no clear connection to sulfur metabolism. This supports previous research into the sulfite metabolism by Borodina *et al.* (2002) suggesting that *H. sulfonivorans* S1 lacks a sulfite oxidase.

Previous research by Koch and Dahl (2018) into sulfur metabolism in another *Hyphomicrobium* species, *H. denitrificans*, also failed to identify a sulfite oxidase for that organism. In response, they proposed that sulfite oxidation to sulfate may occur indirectly via the spontaneous reaction of sulfite and hydrogen sulfide to create thiosulfate (Heunisch, 1977 via Koch and Dahl, 2018), for subsequent oxidation to sulfate via the Sox pathway. It is

therefore possible that a similar mechanism may occur in *H. sulfonivorans* instead of, or in conjunction with, sulfite oxidation via the APS reductase pathway, as proposed by Borodina *et al.* (2002).

Sulfide oxidation

Running counter to sulfite reduction is the more complex process of sulfite oxidation, mediated by the sequential activity of three enzymes: sulfide:quinone oxidoreductase (SQR), rhodanese and persulfide dioxygenase (PDO) (Xia *et al.*, 2017). This begins with the quinone-dependent oxidation of sulfide to produce polysulfide and quinol by sulfide:quinone oxidoreductase. A rhodanese then mediates the reaction of this polysulfide with glutathione (GSH) to generate glutathione persulfide (GSSH), which is then oxidised back to glutathione by persulfide dioxygenase with oxygen and water to yield sulfite.

The sequences of each of these enzymes have already been identified by Koch and Dahl (2018) in *H. denitrificans* ATCC 51888, and a BLAST search for these enzymes in the *H. sulfonivorans* genome suggests that they are also present in this organism. The Pdo-type persulfide dioxygenase has been the most easily identified as C6Y62_00305, showing 66% sequence identify to the PDO of *H. denitrificans*. The search for the Sqr-type sulfide:quinone oxidoreductase and Rhd-type rhadanese suggests that *H. sulfonivorans* contains an Rhd-Sqr fusion protein, C6Y62_00310, instead of two separate enzymes. C6Y62_00310 shows 52% and 44% homology to the *H. denitrificans* Sqr and Rhd respectively, with a C-terminal Rhd domain and N-terminal Sqr domain. Together the presence of these enzymes suggests that *H. sulfonivorans* does indeed contain a pathway for the oxidation of sulfide to sulfite.

Finally, a BLAST and KEGG orthology search (K17229) of the *H. sulfonivorans* genome has failed to identify a strong candidate for an FccAB-type sulfide dehydrogenase associated with the oxidation of sulfide to sulfur in heterotrophic bacteria (Lu *et al*., 2017); sulfur being another potential intermediate of thiosulfate oxidation. Due to the absence of this enzyme and potential redundancy of this enzyme in spontaneous thiosulfate oxidation, the enzyme has been omitted from Figure 3.4-4 for simplicity.

3.4.5 Conclusions

The phenotyping of *H. sulfonivorans* strains in Section 3.2 indicated that the organism can utilise $DMSO_2$ as a sole carbon source via a DMS monooxygenase, and utilise $DMSO_2$,

DMSO and DMS as a sole sulfur source via some other mechanism. A search for alternative enzymes of MSC metabolism in the *H. sulfonivorans* genome has yielded several candidates for two SfnG-like $DMSO_2$ monooxygenases and one MsuD-like MSA monooxygenase. This raises the possibility that *H. sulfonivorans* may have access to a $DMSO_2$ oxidation pathway, which would explain how the organism can still utilise MSCs as a sole sulfur source in the absence of a DMS monooxygenase. However, as both the SfnG and MsuD-like monooxygenases are predicted to generate the methylotrophic intermediate formaldehyde, this does not explain why such a pathway would be specific for assimilatory carbon metabolism rather than assimilatory sulfur metabolism.

As a serine cycle methylotroph (Boden *et al.*, 2011), it is thought that *H. sulfonivorans* assimilates $DMSO_2$ via the generation and metabolism of formaldehyde via the serine cycle. Genomic analysis of formaldehyde assimilation in this organism shows clearly the presence of an H_4MPT pathway, H_4F pathway and the potential for formaldehyde oxidation via two putative formaldehyde dehydrogenases. Based on formaldehyde metabolism in other serine cycle methylotrophs (Crowther *et al*., 2008), a likely mechanism for formaldehyde assimilation may be its oxidation to formate via a combination of the H_4MPT pathway and a formaldehyde dehydrogenase, followed by its conversion to 5,10-methylene H_4F via the H_4F pathway for assimilation via the serine cycle.

Based on the genomic analysis of *H. sulfonivorans* we would expect MSC metabolism to produce two inorganic sulfur compounds, hydrogen sulfide and sulfite, for assimilation as a sulfur source or dissimilation as a by-product of MSC methylotrophy. Various pathways of inorganic sulfur metabolism have been identified in *H. sulfonivorans*, including cysteine synthesis, homocysteine synthesis, sulfate adenylylation, sulfite reduction, sulfite oxidation and a partial Sox system associated with the oxidation of thiosulfate to sulfate. This suggests that inorganic sulfur assimilation is likely to occur via the hydrogen sulfide-dependent synthesis of cysteine and homocysteine (Sekowska *et al*., 2000; Kertesz, 2000), while inorganic sulfur dissimilation is likely to be mediated by via Sox-mediated thiosulfate oxidation or the sulfate adenylylation pathway as previously suggested by Borodina *et al.* (2002).

3.5 Discussion

Although the hypothesis put forward in Section 3.2 that the "DMS monooxygenase pathway is the sole mechanism by which *H. sulfonivorans* can utilise MSCs" was immediately disproven, by the discovery that the *ΔdmoA* strain can still utilise several MSCs as a sole sulfur source, the DmoAB-type DMS monooxygenase still remains a crucial enzyme of methylotrophic MSC metabolism in the organism. The phenotyping of the *H. sulfonivorans ΔdmoA* strain demonstrated that disrupting DmoAB prevents *H. sulfonivorans* from utilising $DMSO_2$ as a sole carbon source, which is consistent with the structure of the DMS oxidation pathway previously discussed by Borodina *et al.* (2000) and Boden *et al.* (2011).

The next logical step would be to update the initial hypothesis to "The DMS monooxygenase pathway is the sole mechanism by which *H. sulfonivorans* can utilise MSCs as a sole carbon source", but this may be a difficult hypothesis to examine without a firmer understanding of the, as yet unknown, alternate mechanism allowing the organism to utilise MSCs as a sole sulfur source in the absence of a functional DmoAB.

An analysis of the *H. sulfonivorans* genome has highlighted a two-enzyme $DMSO_2$ oxidation pathway as a potential candidate for this process, in which a molecule of $DMSO_2$ is oxidised to produce two molecules of formaldehyde and one molecule of sulfite. While this would indeed provide an alternate mechanism for DmoAB-independent sulfur assimilation, the fact that the pathway is also expected to generate the methylotrophic intermediate formaldehyde raises a new issue. If the methylotrophic metabolism of $DMSO_2$ via the alternative $DMSO_2$ oxidation pathway also involves the generation of formaldehyde, then we would also expect a $DMSO_2$ oxidation pathway to facilitate methylotrophic $DMSO_2$ metabolism in the absence of any other extenuating circumstances preventing the pathway's expression or activity.

With this inconsistency in mind, two new hypotheses are proposed below:

The methylotrophic DMS monooxygenase pathway is the sole mechanism by which* H. sulfonivorans *utilises $DMSO_2$ as a sole carbon source.

The alternative $DMSO_2$ oxidation pathway facilitates the utilisation of $DMSO_2$ by* H. sulfonivorans *as a sole sulfur source but not as a sole carbon source.

Chapter 4:

Functional genomics of MSC metabolism in *Hyphomicrobium sulfonivorans* S1

4. Functional genomics of MSC metabolism in *Hyphomicrobium sulfonivorans* S1

4.1 Introduction to the functional genomics of MSC metabolism in *H. sulfonivorans*

In the previous chapter, it was shown the *H. sulfonivorans* is capable of utilising several methylated sulfur compounds (MSCs) as a sole source of carbon and/or sulfur, respectively mediated by a DMS monooxygenase dependent and independent mechanism. Functional analysis of MSC metabolism in the *H. sulfonivorans* genome has generated several putative metabolic pathways for their degradation, dissimilation and assimilation, which can be divided into three broad processes: MSC oxidation, formaldehyde metabolism and inorganic sulfur metabolism. However, more experimental evidence is needed to interrogate the legitimacy of these pathways and inform which specific enzymes in *H. sulfonivorans* genome are involved. This particularly applies to the prospective pathways of $DMSO_2$ metabolism, for which numerous enzymes have been detected experimentally but have yet to be identified (Borodina *et al.*, 2000).

The experiments described in this chapter have applied transcriptomics and proteomics to study putative pathways of MSC metabolism and search for new gene candidates encoding these unidentified enzymes. A comparative transcriptomics experiment has used RNA sequencing to investigate *H. sulfonivorans'* gene expression when the organism utilises $DMSO_2$ as a sole carbon source, while two proteomics experiments have used mass spectrometry to investigate protein expression when *H. sulfonivorans* utilises $DMSO_2$ as a carbon and/or sulfur source. Together, they have been used to construct a multi-omics data set to improve our understanding of the molecular mechanisms *H. sulfonivorans* uses to utilise MSCs.

4.2 Comparative Transcriptomics of $DMSO_2$ Metabolism in *H. sulfonivorans*

4.2.1 Aims and experimental design

The intention of using comparative transcriptomics to study *H. sulfonivorans* was to establish which of the organism's genes and metabolic pathways were associated with the utilisation of $DMSO_2$ as a sole carbon source. The transcriptome data could then be used in

conjunction with the metabolic pathways generated for the organism in Chapter 3, to infer the molecular mechanisms that are involved in $DMSO_2$ utilisation.

Comparative transcriptomics requires two data sets; a control transcriptome which provides baseline expression data for each gene and a test transcriptome in which the organism has been perturbed to produce a test-specific shift in gene expression. In this experiment, the perturbation condition will be the replacement of the control carbon source of MeOH with a carbon source of $DMSO_2$. Genes that are specifically associated with $DMSO_2$ utilisation will then be overrepresented in the test condition transcriptome versus the control transcriptome, while genes that are either unrelated to $DMSO_2$ utilisation or essential for both carbon sources will be will see little shift in gene expression.

The experiment was performed in parallel for the control and test condition in biological triplicate. This can be broken down into four stages (Figure 4.2-1), described below in greater detail for each stage of the process: Cultivation of bacterial biomass, RNA extraction, RNA sequencing and differential gene expression analysis.

Cultivation of H. sulfonivorans biomass

H. sulfonivorans was cultivated in continuous culture under carbon limited conditions, on a sole carbon source of $DMSO_2$ (MSC condition) against MeOH (Control condition), both supplemented with sulfate as an inorganic sulfur source. It was decided that *H. sulfonivorans* would be grown in continuous culture to protect against nutrient depletion within the culture and reduce the indirect effects of MSC metabolism on the transcriptome. For example, when *Hyphomicrobium sulfonivorans* is cultivated on MSCs as a sole carbon source they rapidly acidify their culture media (Borodina *et al.*, 2000), likely due to the excretion of excess sulfur from MSC metabolism into the media as sulfuric acid.

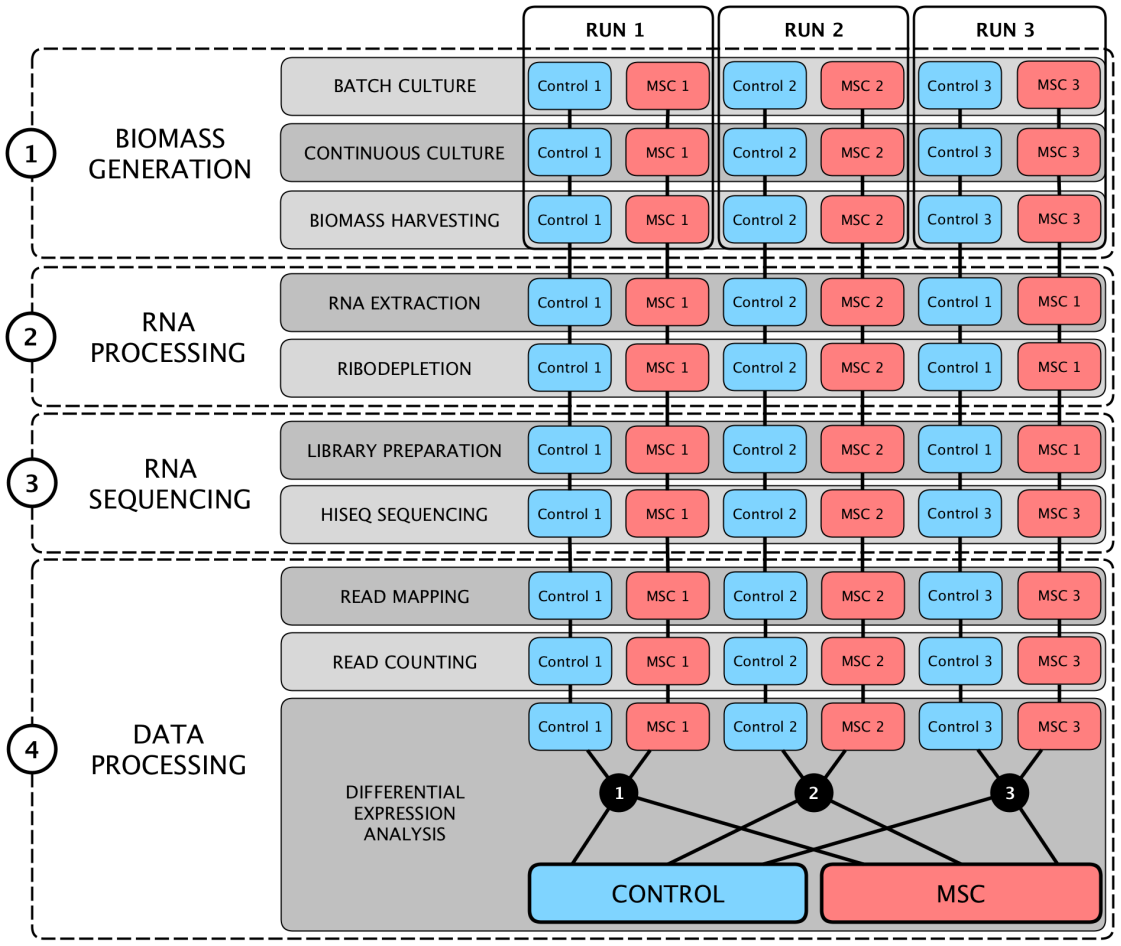


Figure 4.2-1: *H. sulfonivorans transcriptomics experiment workflow*. (1) *H. sulfonivorans* biomass was cultivated under triplicate Control and MSC conditions to yield a total of six biomass samples. Cultivation was performed across three runs in two identical chemostat systems to produce three pairs of Control and MSC biomass. (2) RNA was extracted from each biomass sample, then submitted to Novogene Ltd. for ribodepletion and (3) RNA-sequencing RNA sequencing using an Illumina HiSEQ platform with paired-end reads. (4) Raw reads from the six samples were mapped to the *H. sulfonivorans* S1 genome and the reads counted for each coding sequence. Count data was then used for differential expression analysis of the MSC samples versus the Control samples, normalised pairwise by each chemostat run.

Continuous culture was performed across three runs in two identical chemostat systems, set to provide aeration, agitation and regulate the culture temperature, pH and dissolved oxygen content. Each chemostat consisted of a vessel containing the culture media, an air-line to keep the culture aerated, a flow system to modulate nutrient intake and effluent output, an impeller to keep the culture homogeneous and various mechanisms to regulate the vessel's pH, temperature and dissolved oxygen content (see Section 2.9). Each run, one of the chemostats was used to cultivate *H. sulfonivorans* under the Control condition and the other chemostat used for the MSC condition, alternating between runs to reduce risk of system bias.

Generation of biomass for transcriptomics began with the cultivation of *H. sulfonivorans* in batch culture within the chemostat for 48 h, at which point the chemostat conditions were switched to continuous culture for ~48 h (see Figure 4-2.2). Steady state was achieved in all samples by 18 h of continuous culture and maintained for three volume changes (30 h) to ensure continuity between all samples and conditions. Culture OD_{540} in steady state was consistent between members of the Control and MSC conditions, albeit with the Control showing a higher OD_{540} of ~0.6 compared to the MSC condition's OD_{540} of ~0.2 as is consistent with the cultivation of the organism in batch culture. Bacterial biomass was harvested from continuous cultures at the end of steady state (48-49 h), generating three pairs of Control and MSC condition biomass for transcriptomic analysis.

RNA Extraction, Purification and Sequencing

For each of the triplicate samples of bacterial biomass harvested from the control and test condition cultures, RNA was extracted, purified and checked for RNA quality. The six RNA samples were then submitted to Novogene Ltd. for ribodepletion, library preparation and RNA sequencing. The sequencing for each sample was performed in a single run on an Illumina HiSEQ™ platform, using 150 bp paired-end reads to generate a set of raw read data for each of the six samples for comparative transcriptomic analysis.

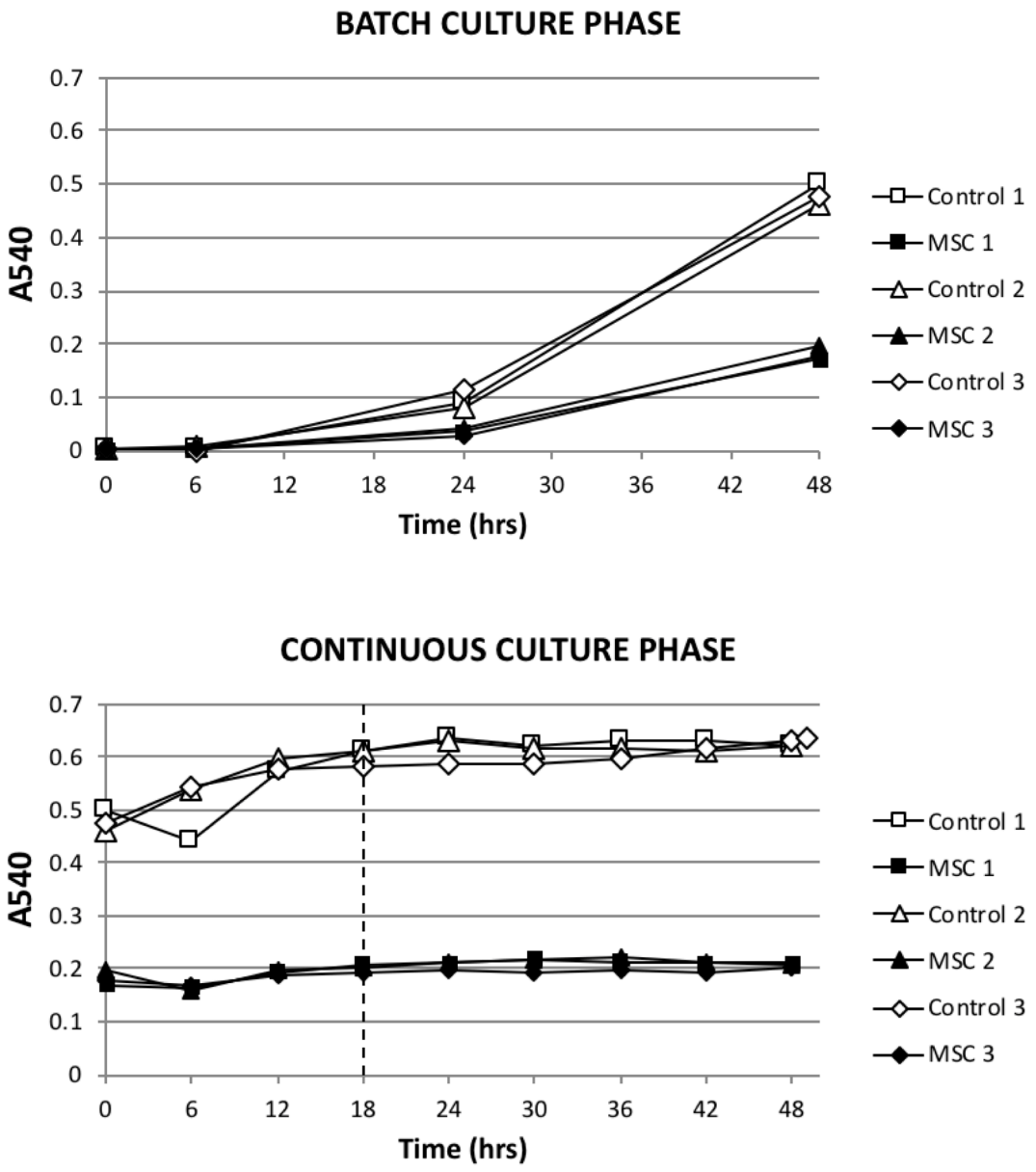


Figure 4.2-2: H. sulfonivorans *chemostat growth curves for each MSC/Control condition pair.* Batch culture phase (TOP) begins with the inoculum of the culture media at 0 h and ends after 48 h growth. In the pre-continuous culture phase (not shown) the chemostat system is calibrated and prepared for continuous culture. In the continuous culture phase (BOTTOM) the system is activated at 0 h, reaching steady state by 18 h (marked by dotted line) and is maintained for three volume changes, ending in biomass extraction at the final time point (48 h for Control 1-2 and MSC 1-3, 49 h for Control 3).

Differential gene expression analysis

Data analysis began with the alignment of paired-end raw read data from each sample to the *H. sulfonivorans* S1 genome. Each sample had ~8-12 million paired-end reads per sample, with an overall alignment rate of over 99% based on concordant and discordant alignment of read pairs (see Figure 4.2-3). Of the concordant read pairs in each sample, in which both members of a pair agree on the alignment site, ~95-96% aligned concordantly once while only ~2-3% aligned concordantly more than once and only 1% of read pairs from each sample aligned discordantly. This high rate of alignment and low rate of multiple alignments suggests that the read data has been successfully mapped to the reference genome. Furthermore, it also suggests that the assembly of the *H. sulfonivorans* S1 2018 draft genome is of sufficiently high quality for a comparative omics analysis.

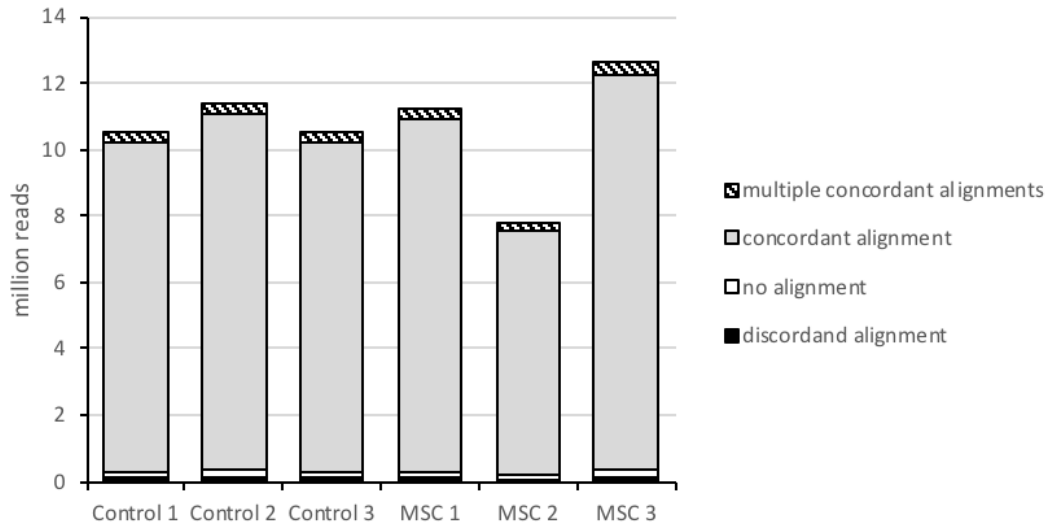


Figure 4.2-3: *Alignment of paired-end raw reads to the* H. sulfonivorans *S1 genome*. Read pairs are categorised as concordant for successful alignment of both members or discordant if each member aligns to a different locus of the genome.

The number of reads aligned to predicted coding sequences in the *H. sulfonivorans* genome was then quantified for each sample to generate three sets of gene count data for each condition. A differential expression analysis of the six samples was performed based on chemostat condition and paired by chemostat run; samples were paired to correct for variance between the runs of continuous culture used to generate bacterial biomass, such as the media batch or inoculum culture used for each run. The normalised differential gene expression data was then used for the transcriptomics of $DSMO_2$ metabolism in *H. sulfonivorans*.

4.2.2 Results overview

The normalised transcriptomics data has been examined using several statistical tests to provide an overview of differential expression in the data set (see Figure 4.2-4), including an MA plot of the data displaying the Log2 fold change (M) for each gene against its base mean (A) across all samples. As the majority of genes in the genome are likely to be constitutively active or repressed across both conditions, we would expect to see that the data as a whole do not show a substantial bias towards either up or downregulation. This plot indicates that the vast majority of the data set centres around a Log2 fold change of 0, i.e. neither up nor downregulated, so appears that the normalisation of the data has been successful and that changes in gene expression are likely to be biological rather than a technical artefact of the sequencing process.

A volcano plot of the transcriptomic data, showing Log2 fold change of each gene against the p-value of this change indicates a promising amount of differentiation in gene expression between the MSC and Control conditions, with a substantial number of genes significantly upregulated (see Figure 4.2-4). This is also the case when examining the data with the more stringent significance measure of adjusted p-value (p-adj), accounting for the likelihood of false positives within the data set. This shows that of the 3,138 genes identified in the transcriptome data, 561 genes show significant upregulation on $DMSO_2$ by p-adj and 116 of these are upregulated by two-fold or more (Log2 fold $\geq$1). Similarly, 548 genes in the data set are downregulated significantly when the organism is cultivated on $DMSO_2$ versus MeOH, with 108 genes downregulated by at least two-fold.

To place this data in the context of MSC compound metabolism, the differential expression data has been mapped onto the various pathways of MSC utilisation that were constructed in Chapter 3: MeOH metabolism, MSC metabolism, formaldehyde metabolism and inorganic sulfur metabolism. A description of differential expression within each pathway, and for other genes of interest, follows below.

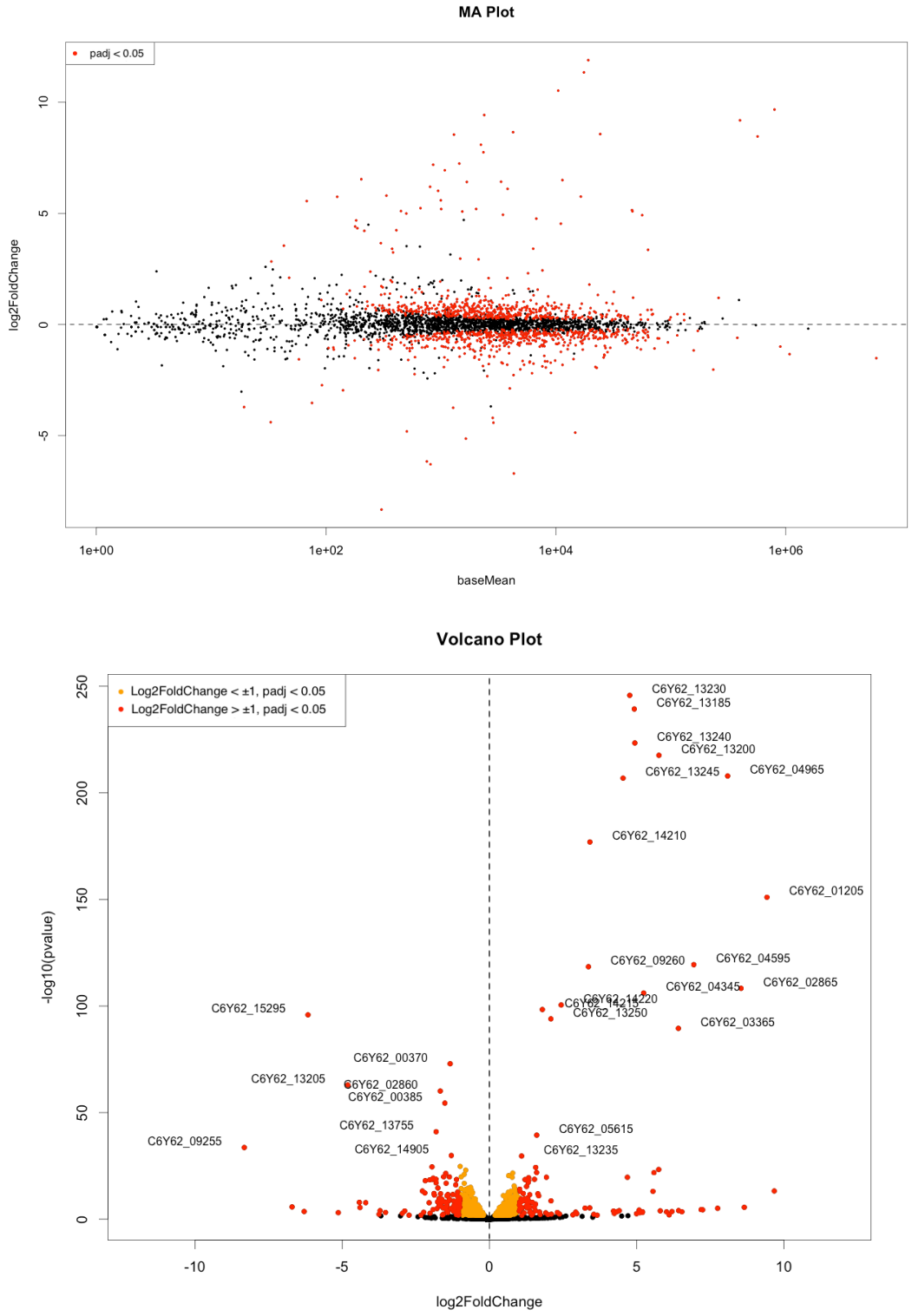


Figure 4.2-4: H. sulfonivorans *transcriptomics data plots.* (TOP) Post-normalisation MA plot for each gene identified in the data set, with Log2 fold-change data expressed as $DMSO_2$/MeOH. Significantly upregulated genes (padj <0.05) are coloured RED. (BOTTOM) *Volcano plot of H. sulfonivorans transcriptome data,* with Log2 fold-change data expressed as $DMSO_2$/MeOH. Significantly upregulated genes (padj <0.05) are coloured ORANGE, significantly upregulated genes that are differentially expressed by >2-fold are coloured RED.

4.2.3 MeOH metabolism

Genome analysis of MeOH oxidation in *H. sulfonivorans* suggested that the organism has an MxaFI-type NADH-dependent MeOH dehydrogenase and a XoxF-type NADH-dependent MeOH dehydrogenase. Transcriptomics shows that each of these genes is expressed by *H. sulfonivorans* when cultivated on a sole carbon source of $DMSO_2$ and MeOH (see Figure 4.2.5). However, in terms of differential expression between the two conditions both *mxaF* (C6Y62_00385) and *mxaI* (C6Y62_00370) are downregulated over 2.5 times when *H. sulfonivorans* was cultivated on $DMSO_2$ versus MeOH, and *xoxF* (C6Y62_12520) is downregulated ~1.5 times.

Looking at other genes of the MxaFJGI MeOH oxidation system, the cytochrome *c* encoding *mxaG* (C6Y62_00375) is downregulated ~1.6 times on $DMSO_2$ versus MeOH, while the protein encoding *mxaJ* (C6Y62_00380) shows no significant change in expression between either condition. However, as with the other genes described above, both *mxaG* and *mxaJ* are also expressed in both conditions.

What this suggests is that expression of MeOH oxidation enzymes are upregulated in the *H. sulfonivorans* transcriptome in response to the use of MeOH compounds as a sole carbon source, but that they appear to be expressed even when *H. sulfonivorans* was cultivated on $DMSO_2$.

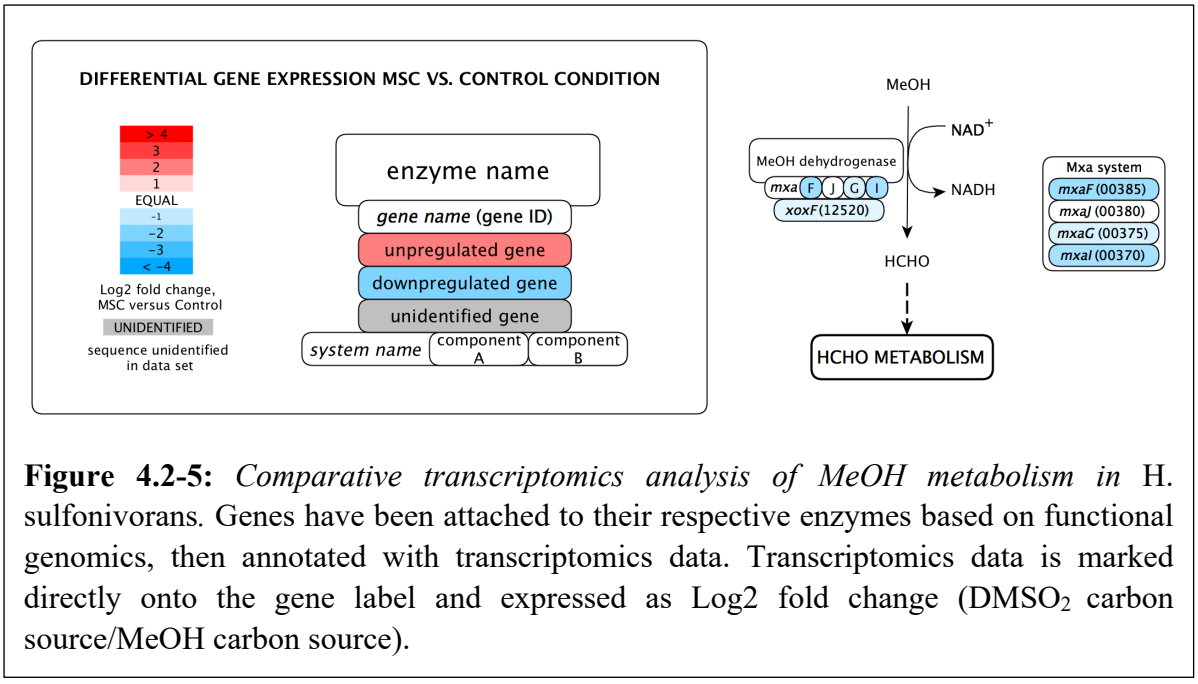


Figure 4.2-5: *Comparative transcriptomics analysis of MeOH metabolism in* H. sulfonivorans. Genes have been attached to their respective enzymes based on functional genomics, then annotated with transcriptomics data. Transcriptomics data is marked directly onto the gene label and expressed as Log2 fold change ($DMSO_2$ carbon source/MeOH carbon source).

4.2.4 MSC Metabolism

The major pathway of MSC metabolism in *H. sulfonivorans* is the DMS monooxygenase pathway, thought to consist of a DMS monooxygenase and an MT oxidase. The *dmoA* gene (C6Y62_13210) encoding the DMS monooxygenase large subunit is expressed on both $DMSO_2$ and MeOH, but is massively upregulated on $DMSO_2$ with an ~800 fold increase in expression. The potential *dmoB* candidate C6Y62_13200 just upstream is also upregulated on $DMSO_2$ by 30-fold. Although a MT oxidase has yet to be identified in *H. sulfonivorans,* the data do show the substantial upregulation of the catalases C6Y62_09260 and C6Y62_15390 when the organism is cultivated on $DMSO_2$. Given that catalases mediate hydrogen peroxide dissimilation and that hydrogen peroxide is a predicted product of MT oxidation, this may also provide indirect evidence for the upregulation of either MT oxidase or another hydrogen peroxide producing enzyme as proposed by Borodina *et al.* (2000).

The alternate $DMSO_2$ oxidation pathway that is predicted to play a role in in MSC metabolism involves a $DMSO_2$ monooxygenase and MSA monooxygenase, for which several candidate genes have been identified. The putative $DMSO_2$ monooxygenase large subunit homologue C6Y62_13190 (*sfnG2*) and small subunit homologue C6Y62_13185 (*sfnF2*) are highly upregulated on $DMSO_2$ compared to MeOH by ~30 fold.

The two other candidates for the large subunits of $DMSO_2$ monooxygenase and MSA monooxygenase are respectively C6Y62_00885 (*sfnG1*) and C6Y62_00835 (*ssuD*), as well as the nearby FMN reductase small subunit C6Y62_00880 (msuE1). However, all three of these genes are only modestly expressed on MeOH and $DMSO_2$, with no significant up or downregulation between the two conditions.

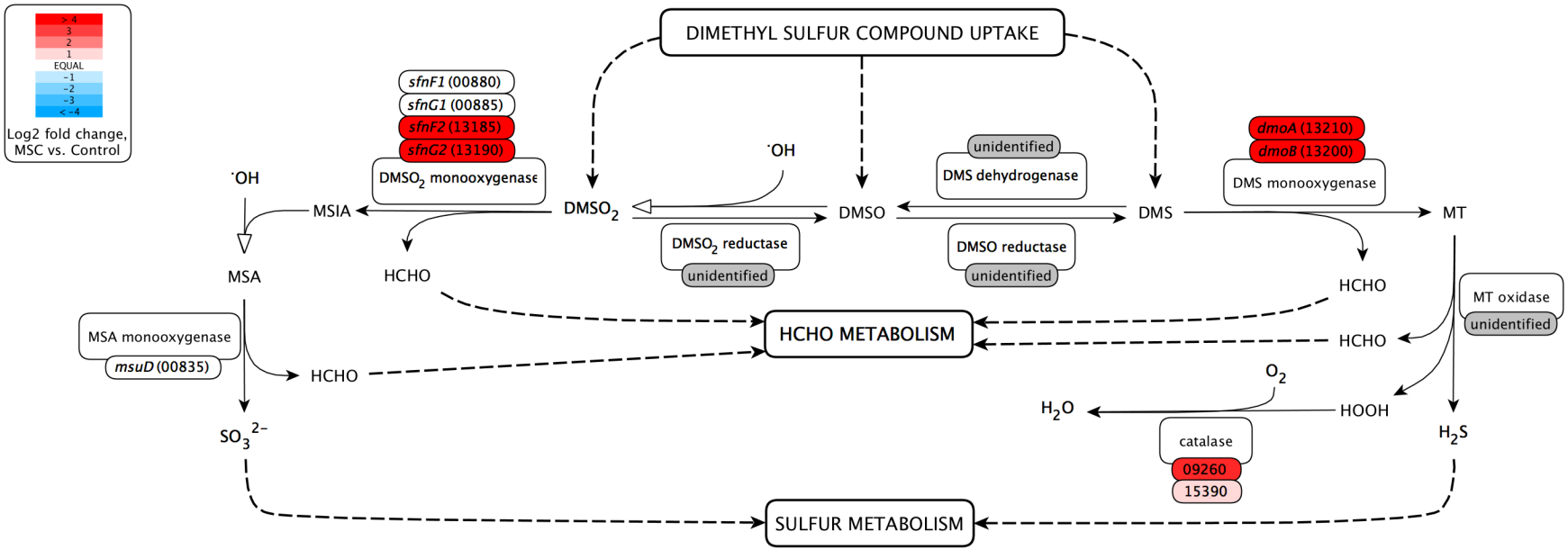


Figure 4.2-6: *Comparative transcriptomics analysis of MSC metabolism in* H. sulfonivorans. Genes have been attached to their respective enzymes based on functional genomics, then annotated with transcriptomics data. Genes with low confidence are marked with a (*) and dotted outline. Transcriptomics data is marked directly onto the gene label and expressed as Log2 fold change ($DMSO_2$ carbon source/MeOH carbon source), with upregulated genes marked in RED, downregulated genes in BLUE and unidentified genes in GREY.

4.2.5 Formaldehyde Metabolism

Although several enzymes of MSC metabolism in *H. sulfonivorans* remain unknown it remains overwhelmingly likely that, as a serine cycle methylotroph, formaldehyde metabolism plays a major role in the organism's assimilation of both MeOH and MSCs as a sole carbon source. The comparative transcriptomics of this process has therefore focused on the four metabolic processes of formaldehyde assimilation predicted by the genome analysis of *H. sulfonivorans*: formaldehyde oxidation, H_4MPT-dependent formaldehyde activation, the H_4F pathway and dissimilatory formate oxidation.

This analysis shows that although the genes encoding the H_4MPT pathway enzymes fae, *fhcABCD, mch* and *mtdB* are all expressed under both conditions, *fae* and *fhcABCD* show a modest but statistically significant upregulation when *H. sulfonivorans* was cultivated on $DMSO_2$ instead of MeOH. Following this trend, the two alternate *fae* homologues C6Y62_05220 and C6Y62_13025 are also upregulated on $DMSO_2$, lending credibility to their potential role as formaldehyde activating enzymes.

In contrast, the putative formaldehyde dehydrogenase encoding gene *fdhA* (C6Y62_10250) and formate dehydrogenase encoding gene *fdh* (C6Y62_02980) are both modestly downregulated in response to the growth of *H. sulfonivorans* on $DMSO_2$. It is possible that this suggests a slight shift in the formaldehyde metabolism on $DMSO_2$ from formaldehyde oxidation to formaldehyde activation via the H_4MPT pathway.

In the H_4F pathway, the *folD* homologue C6Y62_04710 and the *fhs* homologue C6Y62_04850 are expressed when *H. sulfonivorans* was cultivated on both MeOH and $DMSO_2$. This suggests that either H_4F is an important pathway for generating formate and perhaps, as in *Methylobacterium extorquens*, implies that formate is a major intermediate of formaldehyde metabolism via the serine cycle (Crowther *et al.*, 2008). An alternative explanation is that a substantial amount of 5,10-methylene H_4F is being generated by the spontaneous reaction of formaldehyde and H_4F and then dissimilated via the H_4F pathway, but this may be unlikely given that this would be a waste of a valuable carbon source.

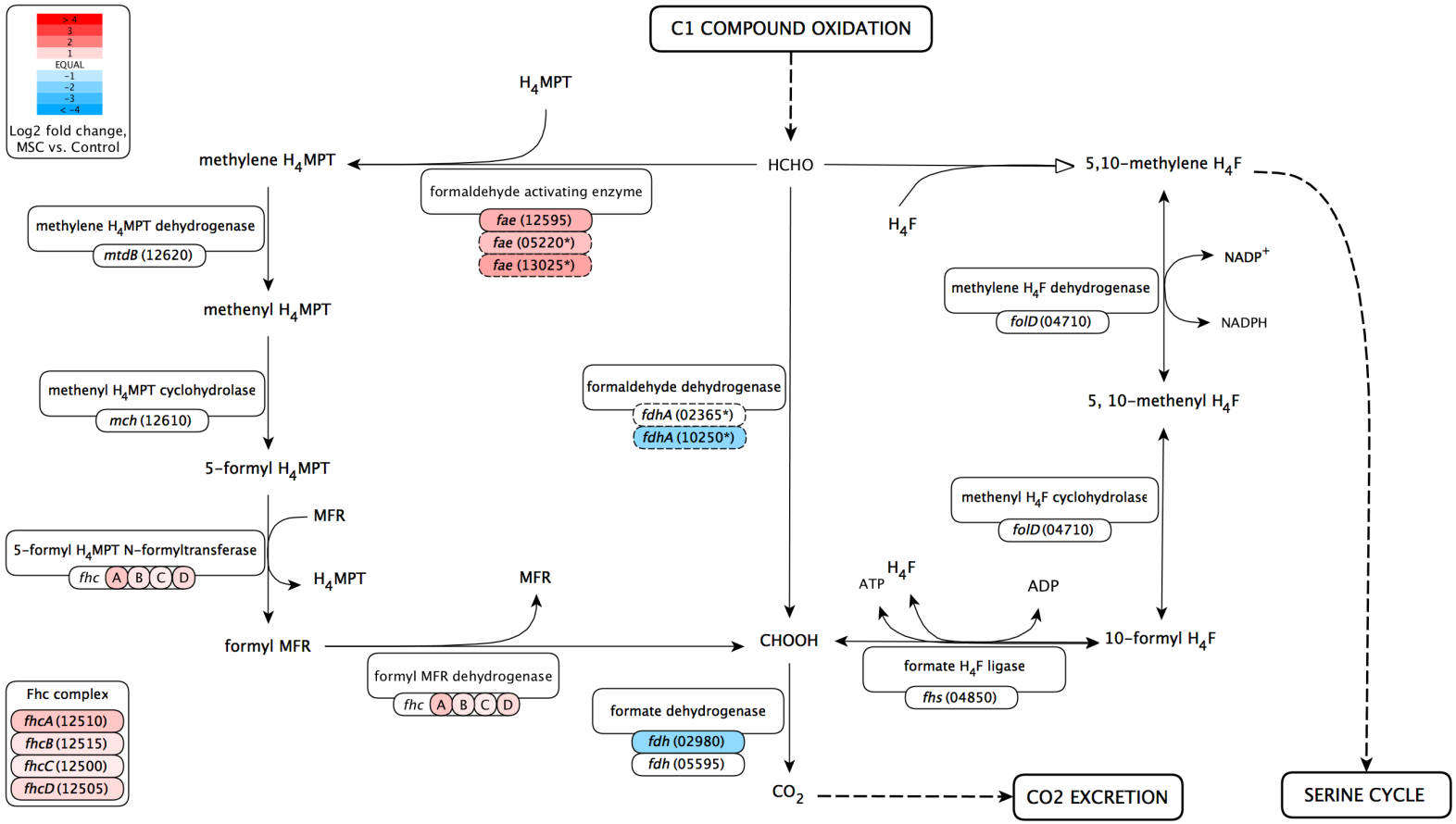


Figure 4.2-7: *Comparative transcriptomics analysis of formaldehyde metabolism in* H. sulfonivorans. Genes have been attached to their respective enzymes based on functional genomics, then annotated with transcriptomics data. Genes with low confidence are marked with a (*) and dotted outline. Transcriptomics data is marked directly onto the gene label and expressed as Log2 fold change ($DMSO_2$ carbon source/MeOH carbon source), with upregulated genes marked in RED, downregulated genes in BLUE and unidentified or unquantified genes in GREY.

In summary, comparative transcriptomics suggests that assimilatory formaldehyde metabolism is likely to occur via the production of formate via the H_4MPT pathway and a putative formaldehyde dehydrogenase, with the formate then fed into the H_4F pathway to generate 5,10-methylene H_4F for carbon assimilation via the serine cycle. Expression data also suggests the presence of formaldehyde dissimilation under both conditions via formate oxidation to carbon dioxide, though it is unknown to what extent this occurs.

4.2.6 Inorganic Sulfur Metabolism

On the MeOH control condition we would expect to see the assimilatory metabolism of sulfate, provided to *H. sulfonivorans* as a sulfur source when cultivated on both the MeOH and $DMSO_2$. In the $DMSO_2$ condition instead we would expect to see the generation of sulfide and/or sulfite as a by-product of methylotrophic $DMSO_2$ metabolism. Although some of this would be utilised as a sole sulfur source, previous research by Borodina *et al.* (2000, 2002) suggests that excess inorganic sulfur is likely to be exported from the cell, likely as sulfate.

Beginning with the reduction of sulfite to hydrogen sulfide, the transcriptomics data shows that the sulfite reductase encoding *cysJ* (C6Y62_12675) is expressed under both conditions with no significant difference between $DMSO_2$ or MeOH. Looking at the oxidation of sulfide to sulfite instead the putative *rhd-sqr* fusion protein (C6Y62_00310) is only moderately downregulated on $DMSO_2$ by ~1.3x and the pdo (C6Y62_00305) shows no significant differential expression.

In the cysteine synthase pathway, which uses hydrogen sulfide to generate cysteine, the putative *cysE* serine O-acetyltransferase C6Y62_09280 and two *cysK* cysteine synthases, C6Y62_15815 and C6Y62_08560, are expressed under both $DMSO_2$ and MeOH. Of these enzymes the *cysK* C6Y62_08560 also shows a small but significant increase in expression in *H. sulfonivorans* cultivated on $DMSO_2$ over MeOH. The expression of this pathway under both conditions implies that cysteine synthesis is not only an important mechanism for assimilating sulfate as a sole sulfur source, but also when *H. sulfonivorans* is metabolising MSCs. This seems consistent with role of cysteine synthesis in other bacterial species as a major mechanism for bacterial sulfur assimilation (Leustek, 2002).

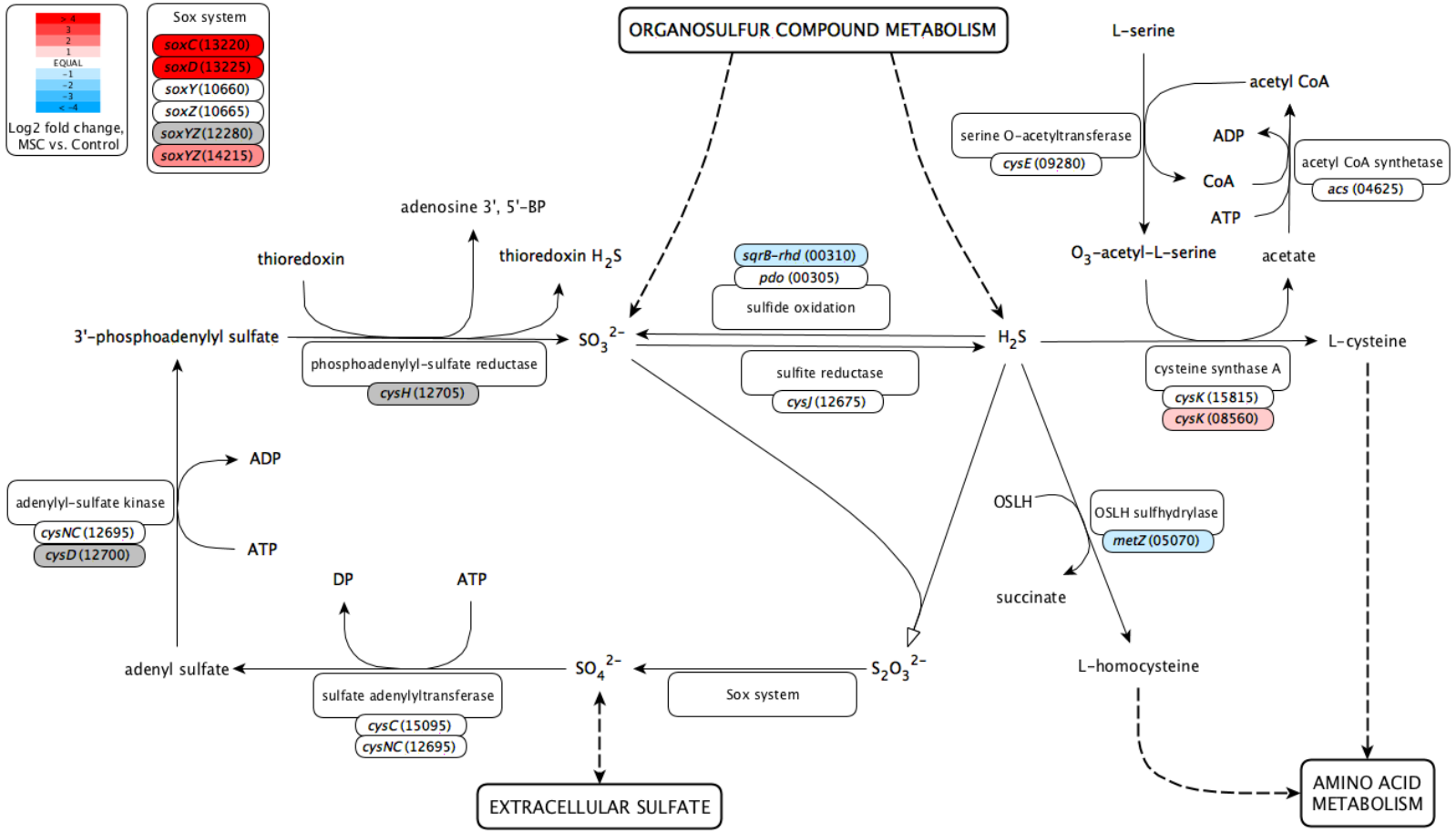


Figure 4.2-8: *Comparative transcriptomics analysis of inorganic sulfur metabolism in* H. sulfonivorans. Genes have been attached to their respective enzymes based on functional genomics, then annotated with transcriptomics data. Genes with low confidence are marked with a (*) and dotted outline. Transcriptomics data is marked directly onto the gene label and expressed as Log2 fold change ($DMSO_2$ carbon source/MeOH carbon source), with upregulated genes marked in RED, downregulated genes in BLUE and unidentified genes in GREY.

Looking at the incomplete Sox system found in *H. sulfonivorans* the *soxC* (C6Y62_13220), *soxD* (C6Y62_13225) and one *soxYZ* (C6Y62_12280) are highly upregulated on $DMSO_2$ versus MeOH but still expressed on both conditions. However, neither the putative *soxY* (C6Y62_14215), *soxZ* (C6Y62_10665) nor other *soxYZ* (C6Y62_14215) genes showed a significant difference in expression between $DMSO_2$ and MeOH, and each exhibited a much lower read count (~10x) than the other Sox genes. Although this suggests that members of the Sox system are playing an active role in the inorganic sulfur metabolism of *H. sulfonivorans* when cultivated in $DMSO_2$, the absence of SoxAXB in the organism's genome and uncertainty surrounding the identities of putative SoxYZ candidates raise questions on the functionality of the Sox system as a whole. Without further information on the function of these genes they have therefore been assigned to the thiosulfate oxidation to sulfate as is the typical function of Sox system proteins (Fredreich *et al.*, 2005).

Although the sulfate adenylation pathway that reduces sulfate to sulfite appears in the *H. sulfonivorans* genome it has not been positively identified in the transcriptomics data set in its entirety, as the phosphoadenylyl-sulfate reductase encoding gene *cysH* and c*ysD* gene encoding the adenylyl-sulfate kinase subunit CysD appear to be missing from the data set. Of the two enzymes encoding genes that are expressed in the data set, *cysC* and *cysNC*, show no significant differential expression $DMSO_2$ versus MeOH. As an inorganic sulfur source of sulfate is present in both conditions, this raises the possibility that *H. sulfonivorans* may be using some other mechanism to generate sulfite from sulfur for use as a sulfur source.

4.2.7 Other upregulated genes of $DMSO_2$ metabolism

A total of 49 genes identified in the *H. sulfonivorans* transcriptome show substantial upregulation greater than 10-fold when cultivated on $DMSO_2$ instead of MeOH, most of which do not map onto the metabolic pathways of MSC metabolism outlined above. A combination of NCBI annotation (Haft *et al.*, 2018) and KEGG orthology (Kanehisa *et al.*, 2016) has been used to infer the roles of these proteins and attempt to match this function back to MSC metabolism. This list of highly upregulated proteins includes several of the enzymes already mapped to the pathways of MSC metabolism discussed above, as well as a potential $DMSO_2$ transporter and several enzymes of cofactor biosynthesis related to the activity of $FMNH_2$-dependent monooxygenases.

To begin with the upregulated enzymes that have already been described in Section 4.2, both pairs of subunits for the DmoAB-type DMS monooxygenase (*dmoA* and *dmoB*) and the

SfnFG-type $DMSO_2$ monooxygenase (*sfnF2* and *sfnG2*) appear on the list as highly upregulated on $DMSO_2$ versus MeOH. The Sox system enzyme SoxCD subunits *soxC* and *soxD* are also present, as is one of the hydrogen peroxide detoxifying catalases.

Located in the same cluster as *dmoA*, *dmoB*, *sfnF2* and *sfnG2* is a porin encoding gene C6Y62_13180, upregulated by ~350 times. A BLAST search for the product of this gene against the Uniprot and NCBI databases reveal no homologues that have been functionally characterised, only that the sequence encodes a porin; porins form transmembrane protein channels for the passive diffusion of molecules across bacterial membranes (Jap & Walian, 1996). As C6Y62_13180 so highly upregulated on $DMSO_2$ and the porin is localised so close to *dmoA*, it raises the possibility that this gene encodes a porin based $DMSO_2$ transporter.

Three of the substantially upregulated genes in the $DMSO_2$ data appear to encode *rib* genes involved in riboflavin biosynthesis. The first is C6Y62_13255, upregulated 50 times and encoding a putative RibH-like 6,7-dimethyl-8-ribityllumazine synthase (K00794) (García-Ramírez *et al.*, 1995). Two more genes, C6Y62_13195 and C6Y62_13230, are upregulated 27 and 35 times respectively and both encode a putative RibBA-like 3,4-dihydroxy 2-butanone 4-phosphate synthase/ GTP cyclohydrolase II (K14652) based on KEGG annotation (Herz *et al.*, 2000). All three of these genes are in close proximity to the FMN-dependent monooxygenases *sfnG* (C6Y62_13190) and *dmoA* (C6Y62_13210) in the *H. sulfonivorans* genome, suggesting that the purpose of this upregulation in riboflavin synthesis is to provide cofactors for these MSC oxidising enzymes.

Looking beyond MSC metabolism, the data suggests an upregulation on $DMSO_2$ of various other proteins relating to cofactor biosynthesis, oxidative stress, DNA repair, transcription, translation and amino acid metabolism, as well as various other hypothetical proteins and enzymes of unknown function. These will be discussed in further detail in Section 4.3 if they are also found to be associated with the metabolism of $DMSO_2$ by comparative proteomics. A full list of these upregulated genes is displayed below in Table 4.2-1.

Table 4.2-1a: *Transcriptomics of* H. sulfonivorans*, most highly induced genes on $DMSO_2$ versus MeOH.* Table displaying the 25 most upregulated genes identified in the transcriptomics data on a sole carbon source of $DMSO_2$ versus the MeOH control condition. Genes that have previously been discussed are highlighted in bold.

Accession No.	gene name	NCBI annotation	strand	position	KEGG ID	fold change versus Control condition
C6Y62_11255	-	30S ribosomal protein S3	Reverse	contig4 654015:654812	K02982	3,805.29
C6Y62_11335	-	50S ribosomal protein L1	Reverse	contig4 673440:674192	K02863	2,601.43
C6Y62_11345	*nusG*	transcription termination/antitermination protein NusG	Reverse	contig4 674903:675439	K02601	1,471.37
C6Y62_13210	***dmoA***	**5%2C10-methylene tetrahydromethanopterin reductase**	**Forward**	**contig5 390172:391614**	**K20938**	**815.81**
C6Y62_01205	-	DUF1134 domain-containing protein	Reverse	contig1 273831:274784		687.27
C6Y62_13175	-	hypothetical protein	Forward	contig5 382414:383469		583.27
C6Y62_13810	*pal*	peptidoglycan-associated lipoprotein Pal	Forward	contig6 78769:79278	K03640	402.87
C6Y62_13220	***soxC***	**sulfite dehydrogenase**	**Forward**	**contig5 392325:393596**	**K17225**	**379.97**
C6Y62_02865	-	hypothetical protein	Forward	contig1 632242:632514		373.59
C6Y62_13180	-	porin	Forward	contig5 383678:384979		352.49
C6Y62_04965	-	two-component sensor histidine kinase	Reverse	contig1 1074640:1075902	K15011	272.40
C6Y62_01165	-	formate dehydrogenase accessory sulfurtransferase FdhD	Forward	contig1 261573:262346	K02379	215.13
C6Y62_06250	-	hypothetical protein	Forward	contig2 244023:244811	K03589	151.31
C6Y62_09975	-	GlsB/YeaQ/YmgE family stress response membrane protein	Reverse	contig4 366228:366500		145.86
C6Y62_04595	-	hydroxymethylbilane synthase	Reverse	contig1 991731:992660	K01749	122.69
C6Y62_07635	-	DUF159 family protein	Forward	contig2 553254:553943		92.85
C6Y62_13225	***soxD***	**cytochrome *c***	**Forward**	**contig5 393589:394167**	**K22622**	**90.47**
C6Y62_02050	-	hypothetical protein	Forward	contig1 456531:458135		85.80
C6Y62_03365	-	disulfide bond formation protein B	Reverse	contig1 734913:735434		85.33
C6Y62_10455	-	DNA helicase	Forward	contig4 472991:473695		73.36
C6Y62_08715	-	NADH-quinone oxidoreductase subunit L	Forward	contig4 77351:79294	K00341	68.77
C6Y62_01595	-	hypothetical protein	Reverse	contig1 363545:363979		64.60
C6Y62_09035	-	2Fe-2S ferredoxin	Forward	contig4 159959:160279	K04755	55.70
C6Y62_13200	***dmoB***	**flavin reductase**	**Forward**	**contig5 388235:388765**		**53.98**
C6Y62_03540	*hisH*	imidazole glycerol phosphate synthase subunit HisH	Forward	contig1 768671:769324	K02501	53.71
scale (fold change versus control condition)				1	>1	>10	>100	>1000

Table 4.2-1b: *Transcriptomics of* H. sulfonivorans*, most highly induced genes on $DMSO_2$ versus MeOH (continued).* Table displaying the 50-26th most upregulated genes identified in the transcriptomics data on a sole carbon source of $DMSO_2$ versus the MeOH control condition. Genes that have previously been discussed are highlighted in bold.

Accession No.	gene name	NCBI annotation	strand	position	KEGG ID	fold change versus Control condition
C6Y62_13255	*ribH*	6%2C7-dimethyl-8-ribityllumazine synthase	Forward	contig5 398527:399039	K00794	48.05
C6Y62_12935	-	precorrin-6A synthase (deacetylating)	Forward	contig5 324245:325006	K02228	46.96
C6Y62_04345	-	DUF1751 domain-containing protein	Forward	contig1 932412:933131		37.68
C6Y62_12010	*xth*	exodeoxyribonuclease III	Reverse	contig5 124364:125146	K01142	36.66
C6Y62_11970	-	protein-L-isoaspartate O-methyltransferase	Forward	contig5 111101:111814	K00573	36.62
C6Y62_13195	*ribB*	3%2C4-dihydroxy-2-butanone-4-phosphate synthase	Forward	contig5 387078:388208	K14652	35.43
C6Y62_01955	-	methyltransferase	Reverse	contig1 437406:438245		34.38
C6Y62_13190	***sfnG2***	**dimethyl sulfone monooxygenase SfnG**	**Forward**	**contig5 385844:386962**	**K17228**	**34.11**
C6Y62_12715	-	DUF192 domain-containing protein	Forward	contig5 273943:274443	K09005	33.89
C6Y62_04975	-	SCO family protein	Forward	contig1 1076709:1077308	K07152	31.86
C6Y62_13240	-	acyl dehydratase	Forward	contig5 396128:396577		30.62
C6Y62_13185	***sfnF2***	**FMN reductase**	**Forward**	**contig5 385211:385756**	**K00299**	**30.25**
C6Y62_13230	-	peptide ABC transporter substrate-binding protein	Forward	contig5 394185:394805	K14652	27.16
C6Y62_15670	*trmB*	tRNA (guanosine(46)-N7)-methyltransferase TrmB	Forward	contig8 18941:19630	K03439	25.71
C6Y62_13245	-	MFS transporter	Forward	contig5 396627:397856		23.21
C6Y62_00510	-	hypothetical protein	Forward	contig1 114428:114784		21.24
C6Y62_05850	-	hypothetical protein	Forward	contig2 158304:159068	K14998	20.11
C6Y62_09015	-	MFS transporter	Forward	contig4 153060:154478	K05820	18.90
C6Y62_09325	-	phosphoribosyl-AMP cyclohydrolase	Reverse	contig4 226035:226481	K01496	18.54
C6Y62_12260	-	P-II family nitrogen regulator	Forward	contig5 179270:179605		12.62
C6Y62_08345	-	transfer Agent	Forward	contig4 7220:7594		11.68
C6Y62_14210	-	alkylhydroperoxidase	Reverse	contig6 178821:179174		10.64
C6Y62_07450	-	hypothetical protein	Forward	contig2 518848:520296		10.59
C6Y62_09260	**-**	**catalase**	**Forward**	**contig4 214166:215644**	**K03781**	**10.27**
C6Y62_12775	-	TonB-dependent receptor	Forward	contig5 288326:290635	K02014	9.48
Scale (fold change versus control condition)				1	>1	>10	>100	>1000

4.3.8 Conclusions

The comparative transcriptomics of $DMSO_2$ oxidation in *H. sulfonivorans* S1 has yielded a large data set covering ~90% of the predicted coding sequences of the *H. sulfonivorans* genome. A small but significant group of these genes show substantial differential expression between $DMSO_2$ and MeOH, several of which can be traced to either MSC metabolism, the *dmoA* gene cluster and/or the biosynthesis of cofactors for DmoAB.

Although many of the enzymes of MSC metabolism in *H. sulfonivorans* remain unknown, the trends exhibited in differential expression data of the few genes that have already been identified is quite promising. For example, the *dmoA* gene encoding the DMS monooxygenase that previous characterised by Boden *et al.* (2011) is massively upregulated on $DMSO_2$, while the putative MeOH dehydrogenase *mxaF* is substantially upregulated on the MeOH condition compared to $DMSO_2$. These results will be discussed further in Section 4.5.

4.3 Comparative Proteomics of $DMSO_2$ utilisation in *H. sulfonivorans*

4.3.1 Aims and experimental design

In addition to comparative transcriptomics, comparative proteomics was also used to investigate the metabolism of $DMSO_2$ by *H. sulfonivorans* S1. While comparative transcriptomics focused on the use of $DMSO_2$ as a sole carbon source, comparative proteomics has examined $DMSO_2$ utilisation as a carbon source and sulfur source. The intention of using comparative proteomics to study *H. sulfonivorans* is to establish which of the organism's proteins and metabolic pathways are associated with the utilisation of $DMSO_2$ as a carbon source, and how this compares to the use of $DMSO_2$ as a sulfur source. Together with the genomics and transcriptomics already performed in *H. sulfonivorans*, this will be used to further explore the molecular mechanisms that are involved in $DMSO_2$ utilisation.

Proteomic analysis of $DMSO_2$ metabolism in *H. sulfonivorans* has been split across two proteomics experiments that share a common work flow, displayed in Figure 4.3-1. In both experiments, *H. sulfonivorans* S1 was cultivated under control and test conditions to generate bacterial biomass. Protein was extracted from the biomass, purified and processed to produce bacterial peptides which were analysed by mass spectrometry. Raw peptide data was then mapped to the *H. sulfonivorans* S1 2018 draft genome to generate control and test condition proteomes for differential analysis. The first experiment examined $DMSO_2$ as a sole carbon source, while the second examined $DMSO_2$ both as a sole sulfur source and as a sole carbon/ sulfur source, as described below.

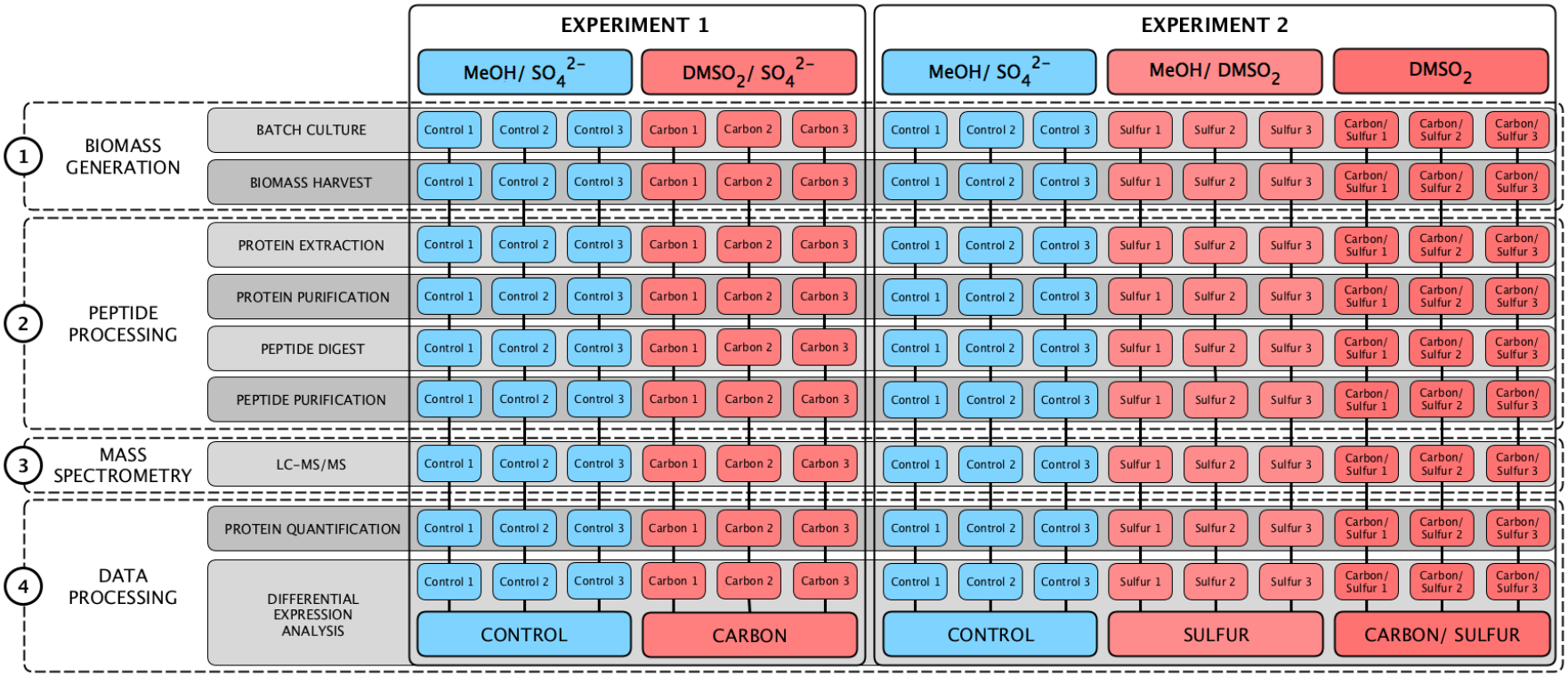


Figure 4.3-1: H. sulfonivorans *proteomics workflow.* Proteomics was performed in two separate experiments following the same workflow: Experiment 1 performed with a Control and Carbon condition, and Experiment 2 with a Control, Sulfur and Carbon/Sulfur Condition. (1) *H. sulfonivorans* was cultivated in triplicate batch cultures to yield a triplicate biomass samples for each condition. (2) Protein was extracted from each biomass sample, purified, digested and the resultant peptides purified to yield triplicate peptide samples for each condition. (3) Mass spectrometry performed by the University of Warwick Proteomics Facility on an LC-MS/MS to produce raw mass spec. date. (4) Raw data mapped to the *H. sulfonivorans* S1 2018 draft genome and quantified using MaxQuant in LFQ-intensity. Protein abundance used for differential expression analysis of each of the test conditions relative to the Control condition of each experiment.

Experiment 1: $DMSO_2$ as carbon source

Experiment 1 used comparative proteomics to examine *H. sulfonivorans* S1 when cultivated on $DMSO_2$ as a sole carbon source (Carbon condition: $DMSO_2$/ SO_4^{2-}) against a non-MSC control carbon source of MeOH (Control condition: MeOH/ SO_4^{2-}).). For each condition, *H. sulfonivorans* was cultivated in triplicate batch cultures on defined, carbon limited media supplemented with a replete sulfur source of sodium sulfate (see Figure 4.3-2). Cultures were harvested after 30 h growth for the Control Condition and 36 h growth for the Carbon condition. Note that while cultivation occurred in parallel, the slower growing Carbon cultures were harvested later than the Control cultures to ensure sufficient biomass for mass spectrometry.

Whole cell protein was extracted from each sample, purified by gel electrophoresis and then digested to produce fragmented protein peptides. These peptides were then analysed by Liquid Chromatography-Mass Spectrometry/ Mass Spectrometry (LC-MS/MS) and peptides mapped to the *H. sulfonivorans* S1 2018 draft genome to generate normalised protein abundance data (LFQ-intensity) in triplicate for each condition.

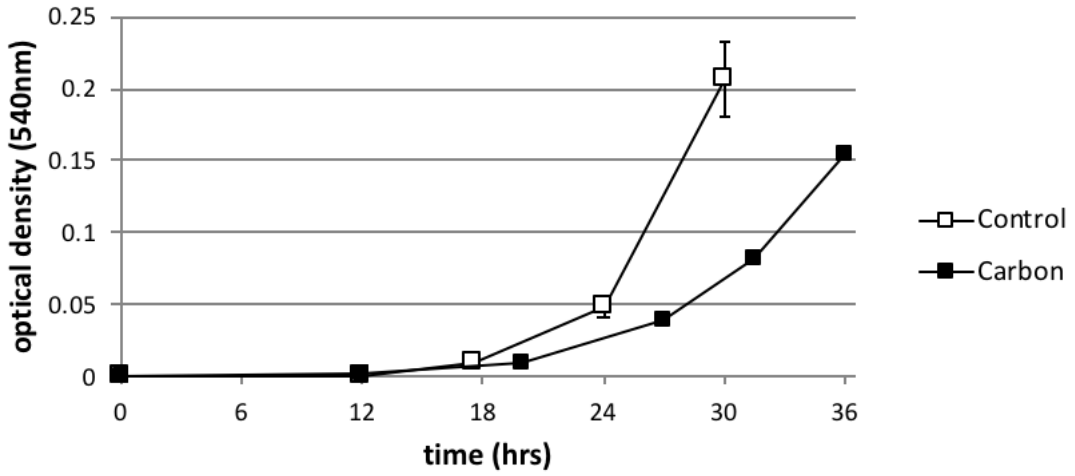


Figure 4.3-2: H. sulfonivorans *proteomics Experiment 1 growth curves. H. sulfonivorans* S1 was cultivated in triplicate batch cultures on a "Control" condition supplemented with MeOH and sulfate, and a "Carbon" condition, supplemented with $DMSO_2$ and sulfate. Biomass was then extracted from cultures at 30 h for the Control cultures and 36 h for the Carbon condition.

Experiment 2: $DMSO_2$ as sulfur and carbon/sulfur source

In the second experiment comparative proteomics was used to examine *H. sulfonivorans* when cultivated on $DMSO_2$ as a sole sulfur source (Sulfur condition: MeOH/ $DMSO_2$) and $DMSO_2$ as a sole carbon and sulfur source (Carbon/Sulfur condition: $DMSO_2$) against non-MSC control substrates (Control condition: MeOH/ SO_4^{2-}). Sulfur condition cultures were supplemented with a non-MSC carbon source of MeOH, while the Control condition cultures contained a sole carbon source of MeOH and a sole sulfur source of sodium sulfate. Note that this was the same control condition used in Experiment 1.

For each condition, *H. sulfonivorans* was cultivated in triplicate batch cultures on defined, carbon limited media and harvested after 30 h growth (see Figure 4.3-3). Whole cell protein was extracted from each sample, purified by gel electrophoresis and then digested to produce fragmented protein peptides. These peptides were then analysed by Liquid Chromatography-Mass Spectrometry/ Mass Spectrometry (LC-MS/MS) and peptides mapped to the *H. sulfonivorans* S1 2018 draft genome to generate normalised protein abundance data (LFQ-intensity) in triplicate for each condition.

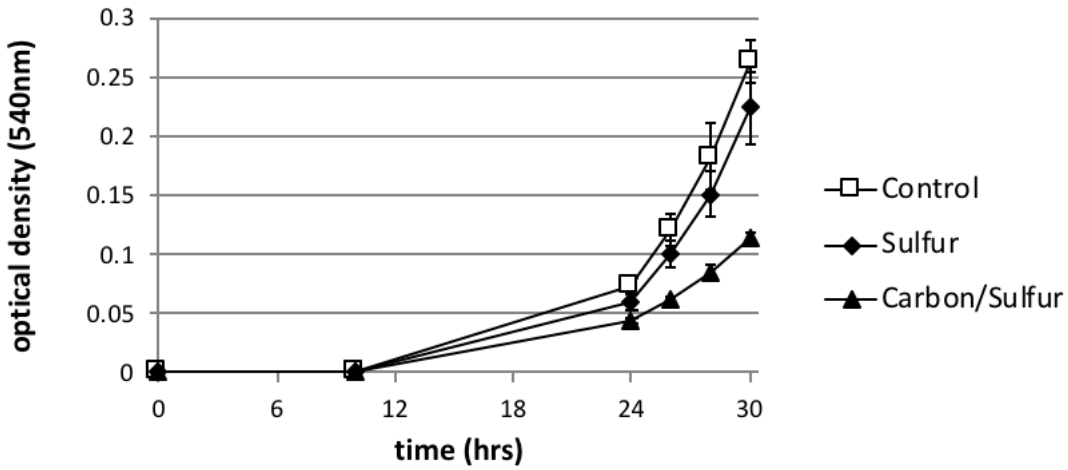


Figure 4.3-3: *H. sulfonivorans proteomics Experiment 2 growth curves. H. sulfonivorans* S1 was cultivated in triplicate batch cultures on either a Control condition (MeOH and sulfate), Sulfur condition (MeOH and $DMSO_2$) or a Carbon/Sulfur condition ($DMSO_2$). Bacterial biomass was harvested from the cultures at 30 h.

4.3.2 Results Overview

The normalised protein abundance data, derived from LFQ-intensity, has been expressed for each test condition relative to the abundance of the MeOH and sulfate control condition. I.e. upregulated proteins are more abundant on the test condition than the control. A total of 1499 protein sequences were identified in Experiment 1, and 1693 proteins were identified in Experiment 2, suggesting that ~40% of the predicted proteome has been identified for each proteomics data set. Note that for a protein to be counted in the data set, it had to be positively identified in three replicates of the same condition, by the presence of at least one unique peptide in the raw mass spectrometry data.

Volcano plots of the proteomics data can be seen in Figure 4.3-4, showing Log2 fold-change of each protein against its -Log p-value. In terms of differential expression each of the $DMSO_2$-containing test conditions shows substantial differential expression compared to their respective Control conditions, indicating a major shift in protein expression when *H. sulfonivorans* was cultivated on $DMSO_2$ as a sole carbon source, sulfur source or carbon and sulfur source. However, in terms of statistical significance (q-value <0.05, false discovery rate 0.05) this shift appears to be far more significant between the Control condition and the Carbon or Carbon/Sulfur conditions than it is in the Sulfur condition.

Note that this doesn't necessarily mean that the differential expression shown by the Sulfur condition is unreliable or insignificant, as several enzymes have been very highly upregulated in the Sulfur condition proteome compared to the Control condition, but that the proportion of proteins showing significant differential expression between the Sulfur and the Control condition is smaller than that of the other conditions. Indeed, this may indicate that the *H. sulfonivorans*' response to the utilisation of $DMSO_2$ as a sole carbon source is greater than its response to $DMSO_2$ as a sole sulfur source.

To place this data in the context of MSC compound metabolism, the proteomics data has been mapped onto the various pathways of MSC utilisation alongside the existing transcriptomics data from Section 4.2. A description of differential expression within each pathway, and for other genes of interest, follows below.

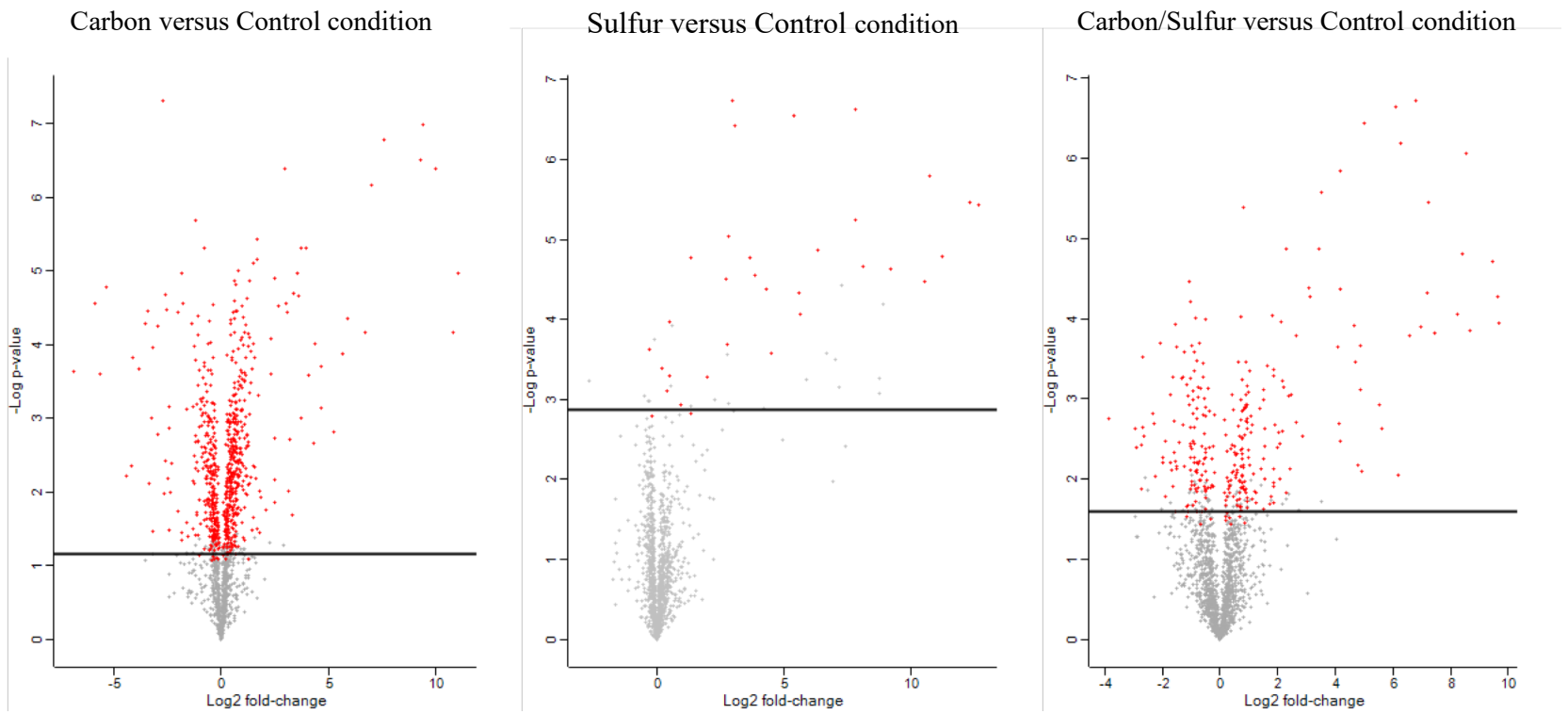


Figure 4.3-4: H. sulfonivorans' *proteomics volcano plots.* (LEFT) Carbon condition ($DMSO_2$ & Na_2SO_4) versus Control condition (MeOH & Na_2SO_4). (MIDDLE) Sulfur condition (MeOH & $DMSO_2$) versus Control condition (MeOH & Na_2SO_4). (RIGHT) Carbon/Sulfur condition ($DMSO_2$) versus Control condition (MeOH & Na_2SO_4). *H. sulfonivorans* proteomics data is expressed as Test Condition/Control Condition. Significantly upregulated proteins are coloured RED (q-value <0.05, FDR 0.05).

4.3.3 MeOH Metabolism

Genomics previously identified two putative NADH-dependent MeOH dehydrogenases in the *H. sulfonivorans* genome, MxaFI and XoxF, both of which were found to be significantly downregulated in the organism's transcriptome when cultivated on $DMSO_2$ in the Carbon condition versus the MeOH control. The downregulation of the MxaFI-type MeOH dehydrogenase subunits MxaF and MxaI was particularly severe.

A search for the MxaFI-type MeOH dehydrogenase in the proteomics data shows that MxaF (C6Y62_00385) is greatly downregulated when the organism is cultivated on a sole carbon source of $DMSO_2$ (Carbon and Carbon/Sulfur conditions) compared to the MeOH containing Control and Sulfur conditions. This is supported by the downregulation of other members of the MxaFJGI cluster on Carbon condition: the dehydrogenase second subunit MxaI (C6Y62_00370), cytochrome *c* MxaG (C6Y62_00375) and electron transfer protein MxaJ (C6Y62_00380).

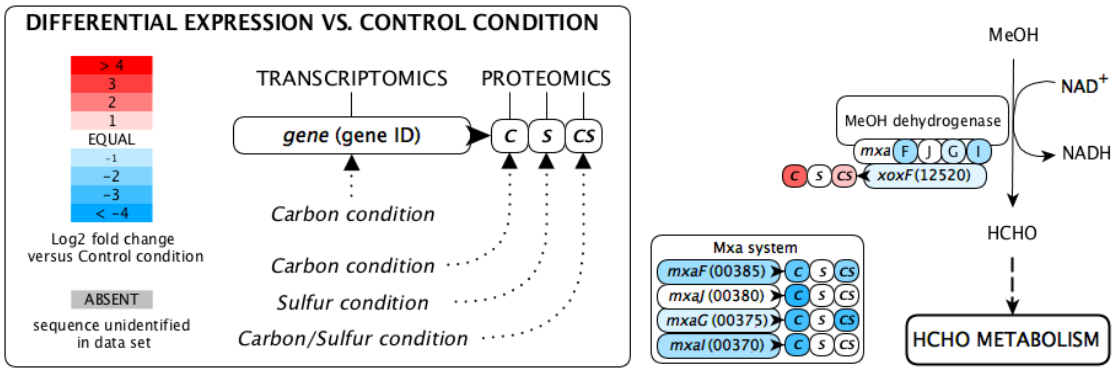


Figure 4.3-5: *Comparative omics analysis of MeOH oxidation in* H. sulfonivorans. General annotation for proteomics data: Gene labels have been attached to their respective enzymes based on the genome analysis performed in Chapter 3, then annotated with proteomics and transcriptomics data. Transcriptomics data is marked directly onto the gene label and expressed as Log2 fold change ($DMSO_2$ carbon source/MeOH carbon source), with upregulated genes marked in RED, downregulated genes in BLUE and unidentified genes in GREY. Proteomics data for the Carbon (C), Sulfur (S) and Carbon/Sulfur (CS) data sets are attached to the gene label expressed as Log2 fold change relative to their respective Control condition, with upregulated proteins marked in RED, downregulated proteins in BLUE and undetected proteins in GREY.

Contrary to the transcriptomics data, the XoxF-like MeOH dehydrogenase (C6Y62_12520) appears to be upregulated in the proteome when cultivated on a carbon source of $DMSO_2$ over MeOH in the Carbon condition (~5 fold) and Carbon/Sulfur condition

(~2 fold). This is somewhat surprising due to the large role played by XoxF-type MeOH dehydrogenases in the MeOH oxidation of other methylotrophic bacteria (Chu *et al.*, 2016), and although it would be premature to suggest that this dehydrogenase plays a direct role in $DMSO_2$ degradation it implies that MxaFI is *H. sulfonivorans*' dominant MeOH oxidising enzyme.

4.3.4 MSC Metabolism

Transcriptomics indicated that when *H. sulfonivorans* was cultivated on $DMSO_2$, genes encoding the DMS monooxygenase *dmoA* (C6Y62_13210) of the DMS oxidation pathway and the putative $DMSO_2$ monooxygenase *sfnG* (C6Y62_13190) of the $DMSO_2$ oxidation pathway are massively upregulated when compared to the MeOH control condition. However, two other potential enzymes of MSC metabolism, the putative MSA monooxygenase C6Y62_00835 and $DMSO_2$ monooxygenase C6Y62_00885, showed no significant differential expression, and neither genomics nor transcriptomics has yielded a clear candidate for an MT oxidase-like enzyme.

In the *H. sulfonivorans* proteomics data set the DMS monooxygenase of the DMS oxidation pathway can be seen as a key enzyme of $DMSO_2$ metabolism, with both the DmoA (C6Y62_13210) and putative DmoB (C6Y62_13200) subunits significantly upregulated in the proteome of the Carbon, Sulfur and Carbon/Sulfur conditions compared to their Control conditions. The DmoA subunit was found to be upregulated ~1000 times on the Carbon condition, ~15 times on Sulfur and ~350 times Carbon/Sulfur, while the DmoB ~700 times on Carbon condition, ~20 times on Sulfur and ~370 times on the Carbon/Sulfur condition. It can be inferred that the DMS monooxygenase plays a much greater role in the utilisation of $DMSO_2$ as a carbon source than as a sulfur source, which is consistent with the capacity of the mutant strain *ΔdmoA* to utilise $DMSO_2$ and other MSCs as a sulfur source.

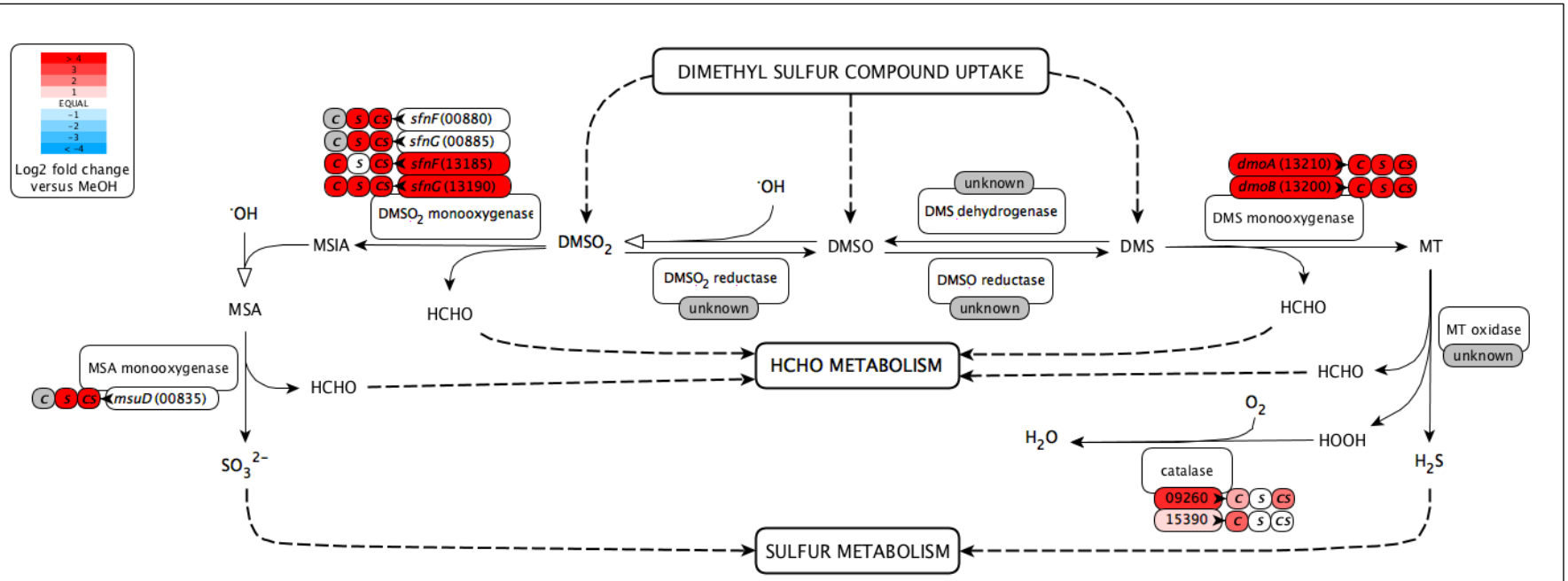


Figure 4.3-6: *Comparative omics analysis of MSC metabolism in* H. sulfonivorans. General annotation for proteomics data: Gene labels have been attached to their respective enzymes based on the genome analysis performed in Chapter 3, then annotated with proteomics and transcriptomics data. Transcriptomics data is marked directly onto the gene label and expressed as Log2 fold change ($DMSO_2$ carbon source/MeOH carbon source), with upregulated genes marked in RED, downregulated genes in BLUE and unidentified genes in GREY. Proteomics data for the Carbon (C), Sulfur (S) and Carbon/Sulfur (CS) data sets are attached to the gene label expressed as Log2 fold change relative to their respective Control condition, with upregulated proteins marked in RED, downregulated proteins in BLUE and undetected proteins in GREY.

In the $DMSO_2$ monooxygenase pathway, the putative $DMSO_2$ monooxygenase subunits SfnG2 (C6Y62_13190) and SfnF2 (C6Y62_13185) that were upregulated in the transcriptome data are also present in the proteomes of all conditions, albeit with some variation in abundance. The SfnG2 subunit was found to be upregulated ~130-fold on the Carbon condition, ~7-fold on Sulfur and ~110-fold Carbon/Sulfur, while the SfnF2 was upregulated by ~290-fold on Carbon condition and ~410-fold on the Carbon/Sulfur condition but showed no significant upregulation on the Sulfur condition. This lower differential expression of SfnFG2 on the Sulfur condition appears to match the protein abundance profile of DmoA, suggesting this monooxygenase also plays a much greater role in the utilisation of $DMSO_2$ as a carbon source than as a sulfur source.

However, the most surprising result comes from the differential expression profile of two other putative enzymes of the $DMSO_2$ oxidation pathway. Here, the SfnFG1-type $DMSO_2$ monooxygenase's SfnG1 (C6Y62_00880) and SfnF1 (C6Y62_00885) subunits were found to be ~2500 times higher under the Sulfur condition and ~300 times higher on the Carbon/Sulfur condition, while seemingly absent from the Carbon condition. Similarly, the MsuD-type MSA monooxygenase (C6Y62_00835) was absent in the Carbon condition but massively upregulated in the Sulfur condition's proteome by ~5000 times and ~700 times higher on the Carbon/Sulfur condition proteome. This suggests that these enzymes were being specifically upregulated in the proteome when *H. sulfonivorans* was cultivated on $DMSO_2$ as a sole sulfur source rather than a sole carbon source.

Indeed, the presence of a sulfur-specific SfnFG-like $DMSO_2$ monooxygenase (C6Y62_00890, C6Y62_00890) and MsuD-like MSA monooxygenase (C6Y62_00835) may provide a mechanism for the DMS monooxygenase-independent metabolism of MSCs and explain how the *ΔdmoA* mutant is still capable of utilising a range of MSCs as a sole sulfur source. This may also free the alternate SfnFG2 (C6Y62_13185, C6Y62_13190), initially presumed to act as a $DMSO_2$ monooxygenase, as a potential candidate for one of the unidentified enzymes of methylotrophic $DMSO_2$ metabolism.

4.3.5 Formaldehyde Metabolism

Formaldehyde metabolism is a key stage of assimilatory C1 compound metabolism in serine cycle methylotrophs, involving the production of 5,10-methylene H_4F for assimilation via the serine cycle. Enzymes of the formate oxidation, formaldehyde oxidation, H_2MPT-dependent formaldehyde activation and H_4F pathways were previously identified in both the

H. sulfonivorans genome and transcriptome data. Many of the putative proteins thought to be encoded by these genes have also been identified in the proteomes of *H. sulfonivorans*, as described below.

Beginning with formaldehyde oxidation, the putative Fdh-like formaldehyde dehydrogenase C6Y62_02365 has been found in the *H. sulfonivorans* proteome under all conditions and is modestly upregulated in the Carbon and Carbon/Sulfur condition. The second putative Fdh-like formaldehyde dehydrogenase, C6Y62_10250, were not identified in any of the proteomes, despite being identified in the transcriptome and being downregulated when $DMSO_2$ was utilised as a carbon source. This may either suggest that these enzymes are only present at low abundance or present in the proteome but not positively identified by mass spectrometry.

The H_4MPT-dependent formaldehyde activation pathway in *H. sulfonivorans* is thought to consist of several Fae formaldehyde activating enzymes (C6Y62_12595, C6Y62_05220 and C6Y62_13025), an MtdB methylene H_4MPT dehydrogenase (C6Y62_12620) and a bifunctional multi-subunit FhcABCD (FhcA C6Y62_12510, FhcB C6Y62_12515, FhcC C6Y62_12500 and FhcD C6Y62_12505) that acts as a 5-formyl H_4MPT N-formyltransferase and formyl MFR dehydrogenase. All of these enzymes have been identified in the *H. sulfonivorans* proteome when cultivated under all conditions, supporting the proposition that this pathway plays an important role in formaldehyde metabolism. In terms of differential expression between proteomes, only the FhcD C6Y62_12505 is modestly upregulated on a sole carbon source of $DMSO_2$ versus MeOH, while the Fae C6Y62_05220 is downregulated. This conflicts with the transcriptome in Section 4.2, which suggested a significant upregulation of the complete *fhcABCD* cluster and all three *fae* genes but is consistent with the lack of differential gene expression shown by the majority of other enzymes in the H_4MPT-dependent formaldehyde activation pathway.

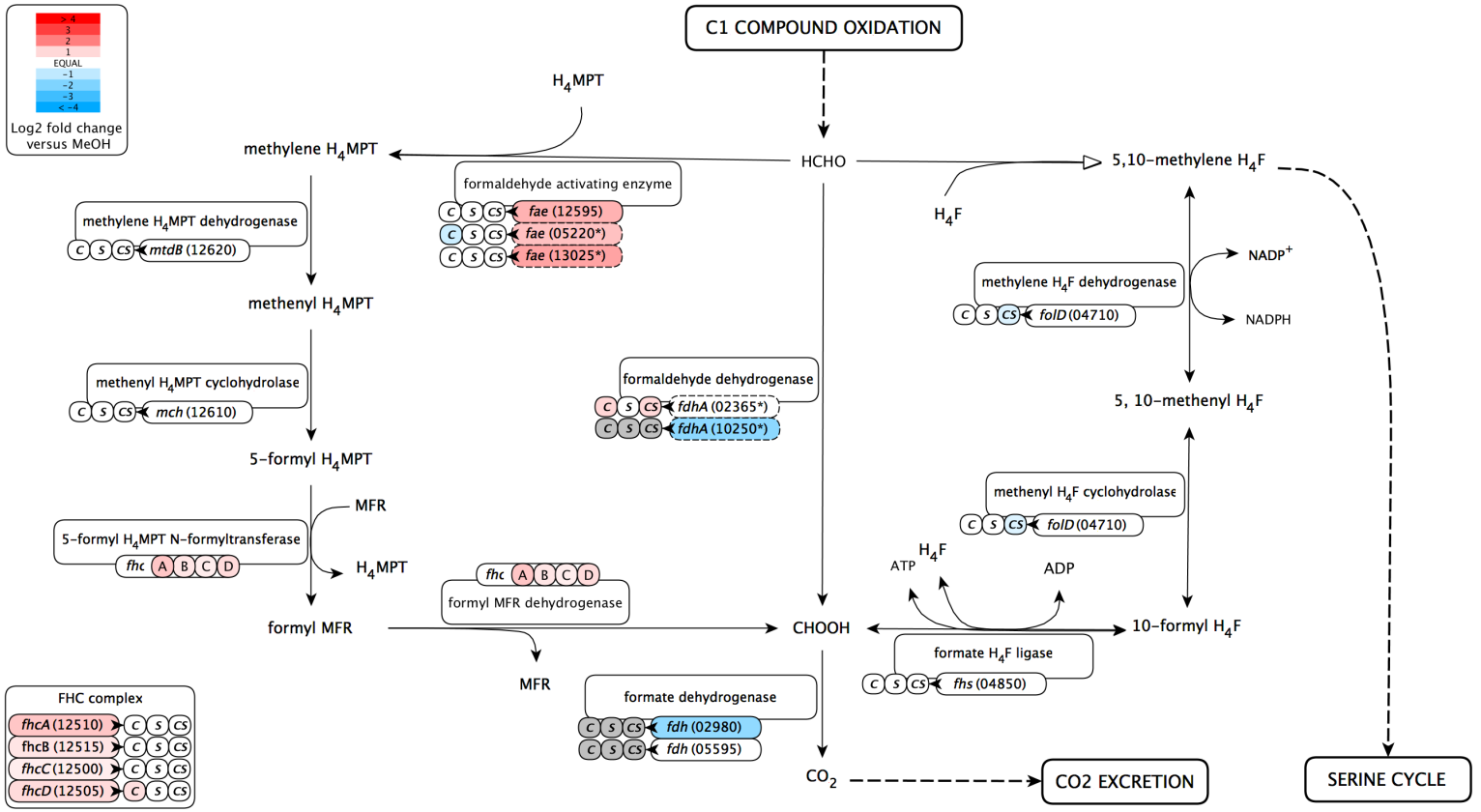


Figure 4.3-7: *Comparative omics analysis of formaldehyde metabolism in* H. sulfonivorans. General annotation for proteomics data: Gene labels have been attached to their respective enzymes based on the genome analysis performed in Chapter 3, then annotated with proteomics and transcriptomics data. Transcriptomics data is marked directly onto the gene label and expressed as Log2 fold change ($DMSO_2$ carbon source/MeOH carbon source), with upregulated genes marked in RED, downregulated genes in BLUE and unidentified genes in GREY. Proteomics data for the Carbon (C), Sulfur (S) and Carbon/Sulfur (CS) data sets are attached to the gene label expressed as Log2 fold change relative to their respective Control condition, with upregulated proteins marked in RED, downregulated proteins in BLUE and undetected proteins in GREY.

Formate dehydrogenases oxidise formate to carbon dioxide, which in *H. sulfonivorans* may represent a mechanism of dissimilatory formaldehyde metabolism. Although two Fae-type formate dehydrogenase homologues (C6Y62_02980 and C6Y62_05595) were successfully identified in the transcriptomics data neither have been identified in any of *H. sulfonivorans* proteome, potentially due to their high similarity making it difficult to positively identify each enzyme.

Finally, the putative H_4F-dependent pathway of *H. sulfonivorans* contains an Fhs methenyl H_4F cyclohydrolase (C6Y62_04850) and a FolD bifunctional formate H_4F ligase/ methylene H_4F dehydrogenase (C6Y62_04710). Both enzymes are expressed in all proteomes, but with a minor downregulation of FolD in the Carbon/Sulfur condition. Given the presence of the H_4MPT pathway and absence of formate dehydrogenase, it may then be possible to infer that these enzymes are most likely to mediate the anabolic metabolism of 5,10-methylene H_4F from formate rather than a mechanism of formate production.

Taken together, this may either support the notion that assimilatory formaldehyde metabolism in *H. sulfonivorans* involves the production of 5,10-methylene H_4F via the H_4F pathway, but may equally suggest that both the H_4F and H_4MPT-dependent formaldehyde activation pathway mediate dissimilatory formaldehyde metabolism.

4.3.6 Inorganic sulfur metabolism

A combination of genomics and transcriptomics have suggested that the assimilatory sulfur metabolism of all four substrate conditions is likely to involve cysteine and/or homocysteine synthesis. The transcriptomics data showed that several Sox enzymes associated with thiosulfate oxidation were highly upregulated in response to the growth of the organism on $DMSO_2$ versus MeOH, but otherwise found that sulfur metabolism expressed only minor differential expression between the different substrates. Due to the complexity of sulfur compound metabolism this description of the proteomics data has been split into the interconversion between sulfite and hydrogen sulfide, hydrogen sulfide assimilation, Sox mediated thiosulfate oxidation and the sulfate adenylylation pathway.

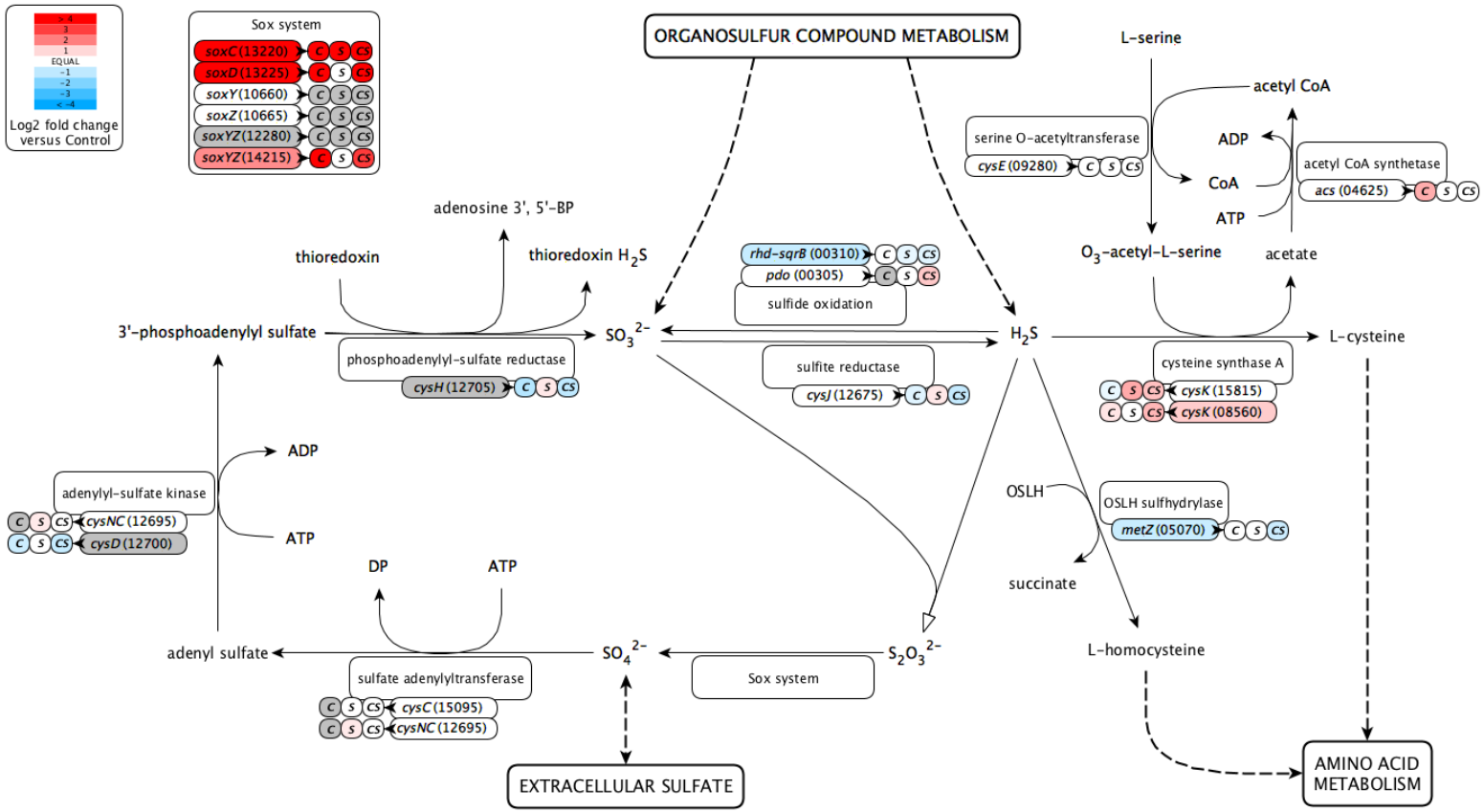


Figure 4.3-8: *Comparative omics analysis of inorganic sulfur metabolism in* H. sulfonivorans. General annotation for proteomics data: gene labels have been attached to their respective enzymes based on the genome analysis performed in Chapter 3, then annotated with proteomics and transcriptomics data. Transcriptomics data is marked directly onto the gene label and expressed as Log2 fold change ($DMSO_2$ carbon source/MeOH carbon source), with upregulated genes marked in RED, downregulated genes in BLUE and unidentified genes in GREY. Proteomics data for the Carbon (C), Sulfur (S) and Carbon/Sulfur (CS) data sets are attached to the gene label expressed as Log2 fold change relative to their respective Control condition, with upregulated proteins marked in RED, downregulated proteins in BLUE and undetected proteins in GREY.

Interconversion of sulfite and hydrogen sulfide

The reduction of sulfite to hydrogen sulfide is predicted to play an important role in the assimilation of both sulfate and sulfite as a sulfur source. The sulfite reductase CysJ (C6Y62_12675) that reduces sulfite to hydrogen sulfide appears to be expressed in the proteome under all carbon and sulfur conditions but is modestly downregulated in the Carbon/Sulfur condition compared to the control condition. As this differential expression is only minor, it suggests that sulfite reduction is an important process for the assimilation of both sulfate and $DMSO_2$ as a sole sulfur source, which is consistent with the proposal above in Section 4.3.5 that the sulfur assimilation of $DMSO_2$ may occur via the sulfite producing $DMSO_2$ oxidation pathway.

Looking at the enzymes responsible for the oxidation of hydrogen sulfide to sulfite, the putative Rhd-Sqr fusion protein (C6Y62_00310) is present in all conditions and only modestly downregulated versus the Control condition when using $DMSO_2$ on the Sulfur and Carbon/Sulfur conditions. The putative Pdo-type persulfide dioxygenase (C6Y62_00305) has not been positively identified in the data set of the first proteomics experiment but is present in all three conditions of the second experiment. However, this Pdo shows no differential expression between the Sulfur and Control conditions, and only modest upregulation on Carbon/Sulfur compared to the Control. Note that as the Pdo was successfully identified in both the transcriptomics data set and the Control condition proteome of Experiment 2, the inability to identify Pdo in Experiment 1 is likely to be a technical artefact of data collection.

These proteomics data sets suggest that the processes of hydrogen sulfide oxidation and sulfite reduction show no substantial changes in their abundance between the various proteomes, perhaps implying that the organism utilises a common mechanism in the metabolism of each set of substrates.

Hydrogen sulfide assimilation

Based on the previous genome analysis, hydrogen sulfide assimilation in *H. sulfonivorans* is likely to be mediated by the synthesis of two sulfur-containing amino acids, cysteine and homocysteine. Cysteine synthesis is a two-step process in which L-serine is acetylated to O_3-acetyl-L-serine, which can then be used with hydrogen sulfide as substrate for cysteine synthesis by a cysteine synthase. The putative CysE-type serine O-acetyltransferase C6Y62_09280 is expressed in all proteomes without significant differential expression. The putative CysK cysteine synthases C6Y62_15815 and C6Y62_08560 are both present in the

proteome under all conditions, but C6Y62_15815 is upregulated in the Carbon and Carbon/Sulfur conditions while C6Y62_08560 is upregulated on the Sulfur and Carbon/Sulfur conditions. This seems to suggest that the process of cysteine synthesis as a whole remains unchanged on the various conditions but that C6Y62_15815 may be somewhat more important for sulfur assimilation and C6Y62_15815 is somewhat more important for $DMSO_2$ utilisation as a sole carbon source, with a shift in abundance barely exceeding ~2.5-fold.

Homocysteine synthesis is a similar process involving the MetZ-type O-succinyl-L-homoserine sulfhydrylase C6Y62_05070, which uses O-succinyl-L-homoserine (OSLH) and hydrogen sulfide to generate succinate and homocysteine for assimilation into the organism's biomass. Although this enzyme is present in all proteomes it shows no differential expression versus MeOH on either the Carbon or Sulfur conditions and only modest downregulation on the Carbon/Sulfur condition. This appears to be consistent with the limited differential expression previously seen in the transcriptomics data and the proteomics of cysteine synthesis above.

Sox-mediated thiosulfate metabolism

The incomplete Sox system in *H. sulfonivorans* appears to contain a SoxC (C6Y62_13220), SoxD (C6Y62_13225), SoxY (C6Y62_10660), SoxZ (C6Y62_10665) and two SoxYZ fusion proteins (C6Y62_12280 and C6Y62_14215), and all except C6Y62_12280 were found in the transcriptome. However, the only Sox proteins to be identified in the proteomics data were the a SoxC (C6Y62_13220), SoxD (C6Y62_13225) and SoxYZ-like carrier protein (C6Y62_14215). Each is upregulated in the *H. sulfonivorans* proteome on sole carbon source of $DMSO_2$ or sole carbon and sulfur source of $DMSO_2$, but only the SoxC C6Y62_13220 is also upregulated on $DMSO_2$ as a sole sulfur source. Based on the proteome and transcriptome data, this suggests that the incomplete Sox system SoxCDYZ is expressed by *H. sulfonivorans* on a sole carbon source of MeOH, but substantially upregulated in response to the growth on $DMSO_2$ as a sole carbon source.

Sulfate adenylylation pathway

The sulfate adenylylation pathway mediates the reduction of sulfate to sulfite, and in *H. sulfonivorans* is thought to contain a CysC sulfate adenylyltransferase (C6Y62_15095), CysH phosphoadenylyl-sulfate reductase (C6Y62_12705) and a bifunctional CysCND sulfate adenylyltransferase/ adenylyl-sulfate kinase consisting of the subunits CysCN (C6Y62_12695) and CysD (C6Y62_126700). All of these are detected in the proteomics data under the Control

condition, Sulfur condition and Carbon/Sulfur condition, while CysD (C6Y62_126700) and CysH (C6Y62_12705) are also present in the proteome of the Carbon condition. It is possible that this failure to identity CysC and CysCN in the Carbon condition is experimental rather biological, as they were also absent from the control of the Carbon condition but present in the control of the Sulfur and Carbon/Sulfur conditions. A closer inspection of CysCN in the proteomics data reveals that although several peptides of this protein were found to be present and downregulated in the Carbon condition, as also seen for CysD and CysH, unfortunately these peptides could not be uniquely assigned to CysCN (C6Y62_12695) as they were shared with the hypothetical protein C6Y62_15955. Taking an overview of the proteomics and transcriptomics data for the adenylyl-sulfate pathway suggests that the pathway is expressed under all carbon and sulfur conditions, with modest downregulation in response to the growth of this organism on $DMSO_2$ as a carbon source.

4.3.7 Other upregulated proteins of $DMSO_2$ metabolism

In addition to the various upregulated proteins discussed above various other proteins are also highly upregulated in the proteomics data in response to the growth of *H. sulfonivorans* on $DMSO_2$. Many of these proteins also show differential expression between the use of $DMSO_2$ as a carbon source and as a sulfur source. The genes that encode the majority of these deferentially expressed proteins are located within two large gene clusters: the *msuD/sfnG1* gene cluster C6Y62_00835:00920 and the *dmoA*/*sfnG2* gene cluster C6Y62_13175:13255.

Carbon sensitive dmoA/sfnG2 *gene cluster* C6Y62_13175:13255

The C6Y62_00835:00930 gene cluster appears to be highly upregulated when $DMSO_2$ is utilised as a sole carbon source, seen in both the Carbon and Carbon/Sulfur conditions, although several members are also upregulated when $DMSO_2$ is utilised as a sole sulfur source (see Table 4.3-1 and Figure 4.3-9). This *dmoA* containing gene cluster has previously been described by Boden *et al.* (2011), so it is interesting to now see the association of the cluster with $DMSO_2$ metabolism highlighted by its differential expression in the proteomics data set. The two most prominent genes of the cluster are the DmoAB-type DMS monooxygenase encoding *dmoA* (C6Y62_13210) and *dmoB* (C6Y62_13200) and the SfnFG-type $DMSO_2$ monooxygenase encoding *sfnF2* (C6Y62_13185) and *sfnG2* (C6Y62_13190), already discussed above in Section 4.3.4. Similarly, the cluster's *soxC* (C6Y62_13220) and *soxD* (C6Y62_13225) genes have been discussed in section 4.3.6.

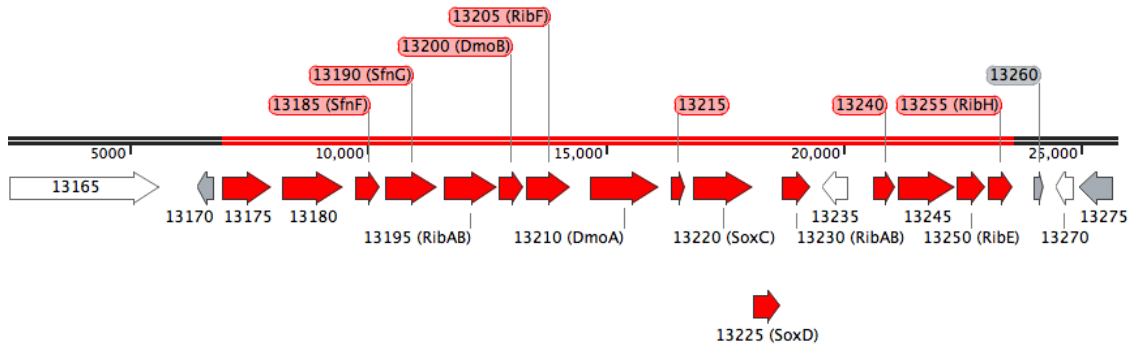


Figure 4.3-9: *H. sulfonivorans dmoA/sfnG2 containing gene cluster C6Y62*_13175:13255. Genes are annotated with their gene ID (prefix C6Y62) and predicted protein product. Each gene is colour coded to represent their expression in the functional genomics data sets (proteomics and/or transcriptomics) when utilising $DMSO_2$ versus a no-MSC Control condition: Upregulated (RED), no significant differential expression (WHITE) and no positive identification (GREY). Sequence length is measured in base pairs and the extent of each gene cluster is marked in onto the DNA in RED.

The cluster also contains several upregulated enzymes relating to riboflavin biosynthesis which are likely to be associated with the FMN for the putative FMN-reductases DmoB (C6Y62_13200) and SfnF2 (C6Y62_13185) which are likely to be important for the activity of their respective $FMNH_2$-dependent MSC monooxygenases. These upregulated enzymes of biosynthesis include two bifunctional RibBA-type 3,4-dihydroxy 2-butanone 4-phosphate synthase/ GTP cyclohydrolase II (C6Y62_13195 and C6Y62_13230) and a

RibF-type riboflavin biosynthesis protein (C6Y62_13205), all three of which were found to be upregulated in the transcriptomics data in Section 4.2.7. Neither the downstream RibE-type riboflavin synthase (C6Y62_13250) and RibH-type 6,7-dimethyl-8-ribityllumazine synthase (C6Y62_13255) have been positively identified in the data set, but as both these enzymes appear to be contiguous to the rest of the cluster and were previously found to be upregulated in the transcriptomics data, they have also been included in the C6Y62_13175:13255 cluster.

The data also shows the upregulation of two membrane transport proteins in this cluster: the upregulated porin previously seen in the transcriptomics data (C6Y62_13180) and a major facilitator superfamily (MFS) transporter (C6Y62_13245). BLAST searches against the UNIPROT database have not found any functionally characterised homologues of either the porin or MFS transporter, so their specific functions are unknown. However, given their proximity to DmoAB and SfnFG2 the transport of extracellular $DMSO_2$ into the cell may be a likely option.

BLAST searches of the third transport protein, annotated as an ABC transporter substrate-binding protein (C6Y62_13230) suggest that this may in fact be a fusion protein (or misannotation) of a GTP cyclohydrolase II (C-terminus) and a RibAB-type Riboflavin biosynthesis protein (N-terminus), which were both previously suggested to participate in Riboflavin biosynthesis by Boden *et al.* (2011) in their examination of the *dmoA* cluster.

One of the most highly upregulated proteins in response to $DMSO_2$ is C6Y62_13215, annotated as an antibiotic biosynthesis monooxygenase but has no known function based on KEGG orthology. Although a BLAST search of the UNIPROT protein database (The UniProt Consortium, 2017) has been unable to identify any experimentally characterised homologues, it did return a sequence from *Hyphomicrobium facile* sp. nov (UNIPROT accession A0A1I7NPV1) annotated as a YgiN-type quinol monooxygenase. This enzyme family has been implicated in quinone redox cycling, where it is thought to oxidise quinol to quinone (Adams and Jia, 2005), but it is uncertanin how such a process would be related to MSC metabolism.

Table 4.3-1: *Proteomics heat map of the* H. sulfonivorans dmoA/sfnG2 *gene cluster C6Y62_13175:13255, plus flanking genes*. Enzymes successfully mapped to the metabolic pathways of MSC utilisation are highlighted in bold.

Accession No.	gene name	NCBI/KEGG annotation	KEGG ID	strand	position (Contig4)	Fold change versus Control condition		
						Carbon: $DMSO_2/SO_4^{2-}$	Sulfur: $MeOH/DMSO_2$	Carbon & Sulfur: $DMSO_2$
C6Y62_13160	-	efflux transporter periplasmic adaptor subunit	K03585	Forward	376537:377946	1.00	1.00	1.00
C6Y62_13165	-	hydrophobe/amphiphile efflux-1 family RND transporter		Forward	377962:381129	2.90	1.00	1.00
C6Y62_13170	-	hypothetical protein		Reverse	381892:382230	-	-	-
C6Y62_13175	-	hypothetical protein		Forward	382414:383469	71.48	-	-
C6Y62_13180	-	porin		Forward	383678:384979	33.99	1.00	48.61
C6Y62_13185	***sfnF2***	**FMN reductase**	K00299	**Forward**	**385211:385756**	**289.09**	**1.00**	**410.18**
C6Y62_13190	***sfnG2***	**dimethyl sulfone monooxygenase SfnG**	K17228	**Forward**	**385844:386962**	**130.84**	**7.02**	**111.00**
C6Y62_13195	*ribAB*	bifunctional RibBA-type 3,4-dihydroxy 2-butanone 4-phosphate synthase/ GTP cyclohydrolase II	K14652	Forward	387078:388208	1,840.59	6.57	68.94
C6Y62_13200	***dmoB***	**flavin reductase**		**Forward**	**388235:388765**	**702.10**	**19.74**	**369.11**
C6Y62_13205	*ribF*	riboflavin biosynthesis protein RibF	K11753	Forward	388821:389762	26.81	1.00	25.83
C6Y62_13210	***dmoA***	**dimethyl sulfide monooxygenase large subunit**	K20938	**Forward**	**390172:391614**	**1,013.41**	**14.51**	**149.61**
C6Y62_13215	-	antibiotic biosynthesis monooxygenase		Forward	391857:392159	941.36	22.51	345.46
C6Y62_13220	***soxC***	**sulfite dehydrogenase**	K17225	**Forward**	**392325:393596**	**782.34**	**42.29**	**808.38**
C6Y62_13225	***soxD***	**cytochrome *c***	K22622	**Forward**	**393589:394167**	**24.16**	**1.00**	**27.79**
C6Y62_13230	*ribAB*	bifunctional RibBA-type 3,4-dihydroxy 2-butanone 4-phosphate synthase/ GTP cyclohydrolase II	K14652	Forward	394185:394805	51.58	5.96	77.87
C6Y62_13235	-	TetR/AcrR family transcriptional regulator		Reverse	395041:395574	0.75	1.00	1.00
C6Y62_13240	-	acyl dehydratase		Forward	396128:396577	7.53	8.39	18.02
C6Y62_13245	-	MFS transporter		Forward	396627:397856	18.29	4.94	18.31
C6Y62_13250	*ribE*	riboflavin synthase	K00793	Forward	397862:398470	-	-	-
C6Y62_13255	*ribH*	6,7-dimethyl-8-ribityllumazine synthase	K00794	Forward	398527:399039	-	-	-
C6Y62_13260	-	hypothetical protein		Forward	399491:399712	-	-	-
C6Y62_13270	-	ETC complex I subunit		Forward	399965:400312	-	-	-
C6Y62_13275	-	hypothetical protein		Reverse	400439:401146	-	-	-
colour scale (fold change versus control condition)			<0.1	<1	1	>1	>10	>100	>1000

Sulfur sensitive msuD/sfnG1 *gene cluster C6Y62_00835:00920*

The C6Y62_00835:00930 gene cluster appears to be highly upregulated when $DMSO_2$ is utilised as a sole sulfur source versus the sulfate containing Control condition, only partially upregulated when $DMSO_2$ is used as a Carbon/Sulfur source but repressed on the Control and Carbon conditions where sulfate is provided as an inorganic sulfur source (see Table 4.3-10). The two most prominent genes of the cluster are the putative SfnFG-type $DMSO_2$ monooxygenase large subunit encoding gene *sfnG1* (C6Y62_00885), MsuDE-type MSA monooxygenase large subunit encoding gene *msuD* (C6Y62_00835) and FMN reductase encoding gene *sfnF1* (C6Y62_00880), which may act as the small subunit of both the SfnFG1 and MsuDE-type monooxygenases. The differential expression of these three proteins has already been described above in Section 4.3.4.

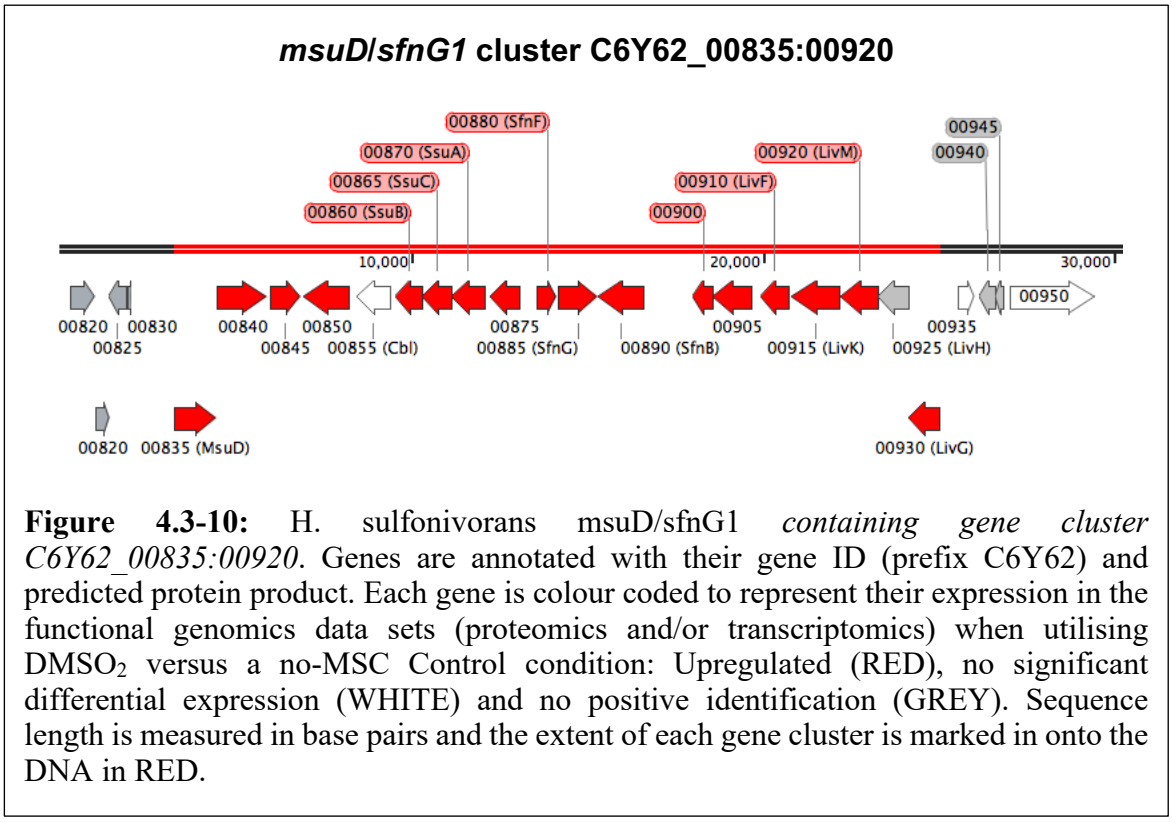


Figure 4.3-10: H. sulfonivorans msuD/sfnG1 *containing gene cluster C6Y62_00835:00920*. Genes are annotated with their gene ID (prefix C6Y62) and predicted protein product. Each gene is colour coded to represent their expression in the functional genomics data sets (proteomics and/or transcriptomics) when utilising $DMSO_2$ versus a no-MSC Control condition: Upregulated (RED), no significant differential expression (WHITE) and no positive identification (GREY). Sequence length is measured in base pairs and the extent of each gene cluster is marked in onto the DNA in RED.

The most highly upregulated protein encoded by this cluster is the cysteine hydrolase (C6Y62_00875), upregulated on the Sulfur condition by >6000-fold compared to the Control condition, although two more hydrolases are also upregulated in the cluster: an amidohydrolase (C6Y62_00845), and a hydroxyacylglutathione hydrolase (C6Y62_00840). Unfortunately, the characterisation of these enzymes is poor with no experimentally characterised close homologues, no predicted function by KEGG orthology and only an approximate NCBI

annotation. What little can be gleaned from this information suggests that they may play a role in the hydrolysis of carbon-sulfur or nitrogen-sulfur bonds.

Like the *dmoA* cluster, the C6Y62_00835:00920 cluster contains several genes encoding putative membrane transport proteins. including another porin of unknown function (C6Y62_00850). The cluster also contains a branched chain-amino acid ABC transport protein, previously identified in *Pseudomonas aeruginosa* (Hoshino *et al.*, 1990), consisting of five subunits: LivF (C6Y62_00910), LivK (C6Y62_00915), LivM (C6Y62_00925), LivH (C6Y62_00925) and LivG (C6Y62_00930). In *Pseudomonas* this system is associated with the membrane transport of branched-chain amino acids, such as L-isoleucine, L-phenylalanine and L-valine, but the relationship between $DMSO_2$ utilisation and these other transporters is currently unclear.

By far the most interesting transport system to be found upregulated in this cluster is an SsuABC-type alkane sulfonate ABC transporter, located just downstream and reverse of the cluster's putative SfnFG1 as three genes: *ssuA* (C6Y62_00870), *ssuB* (C6Y62_00860) and *ssuC* (C6Y62_00865). This transport system was previously characterised in *E. coli* (Eichorn *et al.*, 2000), where it imports alkanesulfonates into the cell for oxidation by an SsuDE-type alkane sulfonate monooxygenase to generate sulfite and an aldehyde. The cluster also contains a putative *cbl* LysR family transcriptional regulator just downstream of *ssuACB*, that has been previously found to regulate the SsuABCDE system of *E. coli* (van der Ploeg *et al.*, 1999). Given the close proximity of this transporter to a putative MsuDE-type methane sulfonate monooxygenase, this SsuABC is a likely candidate for the transport of MSA into *H. sulfonivorans* for assimilatory sulfur metabolism.

Table 4.3-2: *Proteomics heat map of the MsuD/SfnG1 encoding* H. sulfonivorans *gene cluster C6Y62_00835:00930, plus flanking genes*. Enzymes successfully mapped to the metabolic pathways of MSC utilisation are highlighted in bold.

Accession No.	Gene name	NCBI/KEGG annotation	KEGG ID	Strand	Position (Contig1)	fold change versus Control condition		
						Carbon: $DMSO_2/SO_4^{2-}$	Sulfur: $MeOH/DMSO_2$	Carbon & Sulfur: $DMSO_2$
C6Y62_00810	*btuB*	vitamin B12 transporter	K16092	Reverse	177648:179954	0.83	1.00	1.00
C6Y62_00815	-	hypothetical protein		Forward	180717:181442	-	-	-
C6Y62_00820	-	alkane 1-monooxygenase		Forward	181432:181845	-	-	-
C6Y62_00825	-	hypothetical protein		Reverse	181816:182319	-	-	-
C6Y62_00830	-	DUF2497 domain-containing protein		Forward	182369:182479	-	-	-
C6Y62_00835	***msuD***	**alkanesulfonate monooxygenase (Ssud)**	**K04091**	**Forward**	**183693:184880**	**-**	**5,128.90**	**703.15**
C6Y62_00840	-	hydroxyacylglutathione hydrolase		Forward	184877:186322	-	81.83	16.90
C6Y62_00845	-	amidohydrolase	K07045	Forward	186411:187265	-	124.62	1.00
C6Y62_00850	-	porin		Reverse	187347:188642	-	2.91	1.00
C6Y62_00855	*cbl*	LysR family, cys regulon transcriptional activator	K13635	Reverse	188869:189813	-	1.00	1.00
C6Y62_00860	*ssuB*	sulfonate transport system ATP-binding protein	K15555	Reverse	189956:190720	-	30.95	1.00
C6Y62_00865	*ssuC*	sulfonate transport system permease protein	K15554	Reverse	190729:191553	-	4.03	1.00
C6Y62_00870	*ssuA*	sulfonate transport system substrate-binding protein	K15553	Reverse	191556:192515	-	131.45	46.33
C6Y62_00875	-	cysteine hydrolase		Reverse	192673:193491	-	6,435.97	29.31
C6Y62_00880	***sfnF1***	**FMN reductase (SsuE)**	K00299	**Forward**	**193988:194572**	**-**	**224.21**	**32.07**
C6Y62_00885	***sfnG1***	**dimethyl sulfone monooxygenase (SfnG)**	K17228	**Forward**	**194603:195718**	**-**	**277.77**	**29.20**
C6Y62_00890	*sfnB*	SfnB family sulfur acquisition oxidoreductase		Forward	195754:197010	-	2,458.04	175.68
C6Y62_00895	-	hypothetical protein		Reverse	197134:198453	-	18.59	1.00
C6Y62_00900	-	hypothetical protein		Reverse	198443:198997	-	4.66	1.00
C6Y62_00905	-	hypothetical protein		Reverse	199008:200111	-	442.32	1.00
C6Y62_00910	*livF*	ABC transporter ATP-binding protein	K01996	Reverse	200351:201163	-	102.93	1.00
C6Y62_00915	*livK*	ABC transporter substrate binding protein	K01999	Reverse	201228:202598	-	228.57	5.16
C6Y62_00920	*livM*	branched-chain amino acid ABC transporter permease	K01998	Reverse	202635:203678	-	3.75	1.00
C6Y62_00925	*livH*	branched-chain amino acid ABC transporter permease	K01997	Reverse	203682:204578	-	-	-
C6Y62_00930	*livG*	ABC transporter ATP-binding protein	K01995	Reverse	204578:205450	-	145.16	1.00
C6Y62_00935	-	hypothetical protein		Forward	205957:206481	3.28	1.00	1.00
C6Y62_00940	-	DUF305 domain-containing protein		Reverse	206582:207034	-	-	-
colour scale (fold change versus control condition)			<0.1	<1	1	>1	>10	>100	>1000

Sulfate ABC transport gene cluster C6Y62_00835:00920

Another *H. sulfonivorans* gene cluster to show substantial differential expression in the proteomics data is C6Y62_10170:10210, which contains a putative sulfate transport system, hydrolase, porin and a LysrR family transcriptional regulator. This cluster appears to be highly upregulated when $DMSO_2$ is utilised as a sole sulfur source versus the sulfate containing Control condition, only partially upregulated when $DMSO_2$ is used as a Carbon/Sulfur condition but repressed on the Control and Carbon conditions where sulfate is provided as an inorganic sulfur source (see Table 4.3-3).

Unfortunately, the putative porin of this cluster (C6Y62_10210) has not been positively identified in the proteomics data, so it is difficult to assess whether or not this enzyme is upregulated with the *cys* genes directly downstream, nor has it been possible to identify the specific function of this transporter by either homology or KEGG orthology. The same is also true of the putative alpha/beta hydrolase, an enzyme of unknown function, that shows an upregulation of ~4-fold when *H. sulfonivorans* was cultivated on $DMSO_2$ versus sulfate.

As the cluster encodes a putative sulfate transport system that appears to be induced in the absence of sulfate, the most likely role for the cluster is to scavenge extracellular sulfate under low-sulfate availability as this induction of sulfate import has previously been documented in the sulfate limitation response of several other bacterial species (Grabarczyk *et al.*, 2015).

Table 4.3-3: *Proteomics heat map of the sulfate ABC transporter containing* H. sulfonivorans *gene cluster C6Y62_10725:10210, plus flanking genes*.

Accession No.	Gene name	NCBI/KEGG annotation	KEGG ID	Strand	Position (Contig4)	fold change versus Control condition		
						Carbon: $DMSO_2/SO_4^{2-}$	Sulfur: $MeOH/DMSO_2$	Carbon & Sulfur: $DMSO_2$
C6Y62_10155	-	SIMPL domain-containing protein	K09807	Reverse	399518:400228	1.79	1.00	1.00
C6Y62_10165	-	aminoacetone oxidase family FAD-binding enzyme	K07007	Forward	400715:401839	-	0.78	1.00
C6Y62_10170	-	hypothetical protein		Reverse	401959:402582	-	-	-
C6Y62_10175	*cbl*	LysR family transcriptional regulator, cys regulon transcriptional activator	K13635	Reverse	402816:403757	1.00	7.79	4.48
C6Y62_10180	*cysA*	sulfate ABC transporter ATP-binding protein	K02045	Reverse	403754:404839	-	480.05	304.41
C6Y62_10185	*cysW*	sulfate ABC transporter permease subunit	K02047	Reverse	404836:405732	-	8.20	6.32
C6Y62_10190	*cysT*	sulfate ABC transporter permease subunit	K02046	Reverse	405732:406568	-	50.83	17.32
C6Y62_10195	*cysU*	sulfate transport system permease protein		Reverse	406872:408272	-	-	-
C6Y62_10200	-	alpha/beta hydrolase	K07002	Reverse	408416:408985	-	4.19	1.00
C6Y62_10205	*cysP*	sulfate ABC transporter substrate-binding protein	K02048	Reverse	409097:410122	-	1,482.24	124.88
C6Y62_10210	-	porin		Reverse	410373:411758	-	-	-
C6Y62_10215	-	NAD+ synthase	K01916	Reverse	411914:413584	1.00	1.00	1.00
C6Y62_10220	-	transporter	K07088	Forward	413946:414899	-	-	-
C6Y62_10225	-	DUF502 domain-containing protein		Reverse	414954:415733	1.00	1.00	1.00

colour scale (fold change versus control condition)	<0.1	<1	1	>1	>10	>100	>1000

LadA *gene cluster C6Y62_13575:013615*

The fourth and final gene cluster to show substantial differential expression in the proteome data is C6Y62_13575:013615, which contains two monooxygenases and an ABC transport system (see Table 4.3-4). This cluster is highly upregulated when $DMSO_2$ is utilised as a sole sulfur source versus sulfate, only partially upregulated when $DMSO_2$ is used as a Carbon/Sulfur condition but repressed on the Control and Carbon conditions. A search for the cluster in the transcriptomics data also found no differential expression for any of the C6Y62_13575:013615 genes on a sole carbon source of DMSO2 versus a MeOH control condition.

Based on KEGG orthology the transporter appears to be another Liv-type branched-chain amino acid ABC transporter as previously found in the msuD/sfnG1 cluster described above. Although the transporter's LivH (C6Y62_13600) and LivM (C6Y62_13595) subunits are absent from all proteomes, the other subunits LivF (C6Y62_13580), LivK (C6Y62_13585 and C6Y62_13590) and LivG (C6Y62_13605) subunits have been positively identified in the proteomics data from experiment 2, where they are substantially upregulated on the Sulfur condition versus the Control condition.

Also upregulated in this cluster is another $FMNH_2$-dependent monooxygenase, C6Y62_13615, which KEGG orthology suggests may be a LadA-type long-chain alkane monooxygenase. This LadA-type monooxygenase enzyme is apparently related to SsuD, mediates the $FMNH_2$-dependent oxidation of alkanes to alcohols (Feng *et al.*, 2007), and was previously identified as a homologue of DmoA in the genome analysis of *H. sulfonivorans* (see Chapter 3). As no obvious candidates for the enzyme's second FMN reductase subunit have been found either up or downstream of the cluster, it is possible that this enzyme may share the subunit of one of the various other $FMNH_2$-dependent monooxygenases. A point of contention for the function of this putative LadA is that the enzyme was characterised in *Geobacillus* thermodenitrificans from a subsurface oil reservoir, an organism that is far removed from *Hyphomicrobium*, so the true function of this enzyme may be quite different. For the time being C6Y62_13615 will keep the name LadA if only to differentiate it from the other FMNH2-dependent monooxygenases.

Table 4.3-4: *Proteomics heat map of the LadA-encoding* H. sulfonivorans *gene cluster* C6Y62_13575:013615*, plus flanking genes*.

Accession No.	Gene name	NCBI/KEGG annotation	KEGG ID	Strand	Position (Contig4)	fold change versus Control condition		
						Carbon: $DMSO_2/SO_4^{2-}$	Sulfur: $MeOH/DMSO_2$	Carbon & Sulfur: $DMSO_2$
C6Y62_13565	-	redox-regulated ATPase YchF	K06942	Reverse	25391:26491	0.86	1.00	1.00
C6Y62_13570	-	acyl CoA:acetate/3-ketoacid CoA transferase	K01026	Forward	27124:29073	1.00	1.00	1.00
C6Y62_13575	-	monooxygenase		Reverse	29234:30463	-	12.60	1.00
C6Y62_13580	*livF*	ABC transporter ATP-binding protein	K01996	Reverse	30483:31286	-	3.21	1.00
C6Y62_13585	*livK*	ABC transporter permease	K01999	Reverse	31289:32656	-	439.11	11.42
C6Y62_13590	*livK*	ABC transporter permease	K01999	Reverse	32653:33939	-	49.32	2.83
C6Y62_13595	*livM*	branched-chain amino acid ABC transporter permease	K01998	Reverse	34035:35096	-	-	-
C6Y62_13600	*livH*	branched-chain amino acid ABC transporter permease	K01997	Reverse	35108:35989	-	-	-
C6Y62_13605	*livG*	ABC transporter ATP-binding protein	K01995	Reverse	36004:36801	-	3.42	1.00
C6Y62_13610	-	hypothetical protein		Reverse	36798:37772	-	6.85	1.00
C6Y62_13615	*ladA*	long-chain alkane monooxygenase	K20938	Reverse	37777:39180	-	59.26	1.00
C6Y62_13620	-	hypothetical protein		Reverse	39680:42931	-	-	-
C6Y62_13625	-	hypothetical protein		Reverse	43304:43798	-	-	-
C6Y62_13630	-	GAF domain-containing protein		Reverse	43885:44406	-	-	-
C6Y62_13565	-	redox-regulated ATPase YchF	K06942	Reverse	25391:26491	0.86	1.00	1.00
C6Y62_13570	-	acyl CoA:acetate/3-ketoacid CoA transferase	K01026	Forward	27124:29073	1.00	1.00	1.00

colour scale (fold change versus control condition)	<0.1	<1	1	>1	>10	>100	>1000

It is unknown what role is played by a branched-chain amino acid transport system and putative alkane monooxygenase in the utilisation of $DMSO_2$ as a sole sulfur source, but it may be part of the organism's sulfate starvation response. As an alkane monooxygenase, it is possible that the function of the enzyme is to oxidise the long chain substrates or products of SsuDE-like alkanesulfonate monooxygenases, but this is only speculative without further study.

4.3.8 Conclusions

The proteomics data shows strong differential expression in the *H. sulfonivorans* proteomes when cultivated under the various carbon and sulfur conditions, with substantial shifts in protein abundance shown between different carbon sources ($DMSO_2$ and MeOH) and sulfur sources ($DMSO_2$ and sulfate). Although there are some minor instances where the proteomics data conflicts with transcriptomics from Section 4.3, the two data sets are generally in agreement on the differential expression of the various genes and proteins of *H. sulfonivorans*. Indeed, the proteomics data has even provided expression data for several genes that could not be positively identified by the transcriptomics.

More importantly, many of the differentially expressed proteins have been successfully mapped onto the metabolic pathways of MSC metabolism that were proposed in Chapter 3, indicating that these are at least an approximate representation of $DMSO_2$ metabolism by the organism. The majority of these differentially expressed proteins appear to be encoded by four gene clusters, one of which appeared to be sensitive to the utilisation of $DMSO_2$ as a carbon source (C6Y62_13175:13255) and three sensitive to $DMSO_2$ as a sulfur source (C6Y62_00835:00930 and C6Y62_10725:10210, C6Y62_13575:013615).

Comparative proteomics has also highlighted the substantial upregulation of several putative $FMNH_2$-dependent MSC monooxygenases in response to the growth of *H. sulfonivorans* on $DMSO_2$ compared to a non-MSC control condition. What makes this particularly interesting is that these enzymes also show differential expression between the use of $DMSO_2$ as a carbon source and sulfur source, suggesting that several of these enzymes may be involved in the non-methylotrophic metabolism of MSCs, and suggests that these enzymes may be worth further investigation.

4.4 Phylogenetics of $FMNH_2$-dependent monooxygenases

Bacterial $FMNH_2$-dependent monooxygenases have previously been shown to metabolise MSCs in several bacterial species, including the DmoAB-type DMS monooxygenase of *H. sulfonivorans* (Boden *et al.*, 2011), MsuDE-type MSA monooxygenase of *Pseudomonas aeruginosa* (Kertesz *et al.*, 1999), SfnFG-type $DMSO_2$ monooxygenase of *Pseudomonas fluorescens* (Wicht, 2016). Although identifying the function of each $FMNH_2$-dependent monooxygenase will require their purification and experimental characterisation, it may be possible to approximate their roles by categorising them into phylogenetic groups.

To achieve this, a sequence list was constructed containing the $FMNH_2$-dependent monooxygenases from *H. sulfonivorans*, experimentally characterised $FMNH_2$-dependent MSC monooxygenases and related but uncharacterised bacterial $FMNH_2$-dependent monooxygenases. Protein BLAST searches were performed against the UNIPROT database using the amino acid sequences of DmoA (C6Y62_13200), MsuD (C6Y62_00835), LadA (C6Y62_13605), SfnG1 (C6Y62_00885) and SfnG2 9C6Y62_13190) from *H. sulfonivorans* to gather a random selection of homologous sequences from other bacterial species. This ranged from the close homologues of other *Hyphomicrobium* species to poor homologues from more distant members of the Alphaproteobacteria.

These sequences were then aligned against the SsuD-type alkanesulfonate monooxygenase from *Escherichia coli* (Eichhorn *et al.*, 1999), MsuD-type MSA monooxygenase from *Pseudomonas aeruginosa* (Kertesz *et al.*, 1999), SfnG-type $DMSO_2$ monooxygenase from *Pseudomonas fluorescens* (Wicht, 2016) and another $FMNH_2$-dependent monooxygenase from *H. sulfonivorans* identified in Chapter 3 (C6Y62_11755) that appears to be a poor homologue of the other $FMNH_2$-dependent monooxygenases such as DmoA and SfnG. Note that this homologue showed no differential expression in either the proteomics or transcriptomics data sets and has no known function based on KEGG orthology. A phylogenetic tree was constructed from the alignment and the confidence assessed using phylogenetic bootstrapping (bootstrap 100).

If the amino acid sequences of these enzymes are sufficiently distinct to predict their function then we would expect to see strong differentiation between the DmoA, MsuD, SsuD and SfnG-type monooxygenases, with a strong association of the various *Hyphomicrobium* sequences to specific enzyme clusters. However, if the *Hyphomicrobium* sequences form a distinct cluster of their own or form a larger, overlapping cluster without clear differentiation between the reference enzymes then they may be too diverse to categorise by bioinformatics.

4.4.1 Results

Phylogenetic analysis shows good segregation of the $FMNH_2$-dependent monooxygenase sequences into four broad clades with a high degree of confidence (≥0.85), named after their respective genes in *H. sulfonivorans*: DmoA, SfnG, MsuD/SsuD and LadA (see Figure 4.4-1). Note that the annotation of these clades is somewhat arbitrary as an indicator of sequence dissimilarity between clades rather than homogeneity within them.

The putative SfnG and MsuD-type monooxygenases from *H. sulfonivorans* appear to cluster with their respective experimentally characterised homologues from *P. putida* and *P. fluorescens*, while the *H. sulfonivorans* DmoA and putative LacA-type monooxygenases are separated into their own distinct clades. However, a closer inspection of the tree suggests that SfnG2 is part of an outgroup from the rest of the SfnG-like sequences, and that the SsuD-like sequences form an outgroup from the rest of the MsuD/SsuD-like sequences. The unknown, low homology monooxygenase subunit C6Y62_11755 is a clear outlier from any of the other monooxygenase sequences, which is to be expected given its low homology to the known MSC monooxygenases.

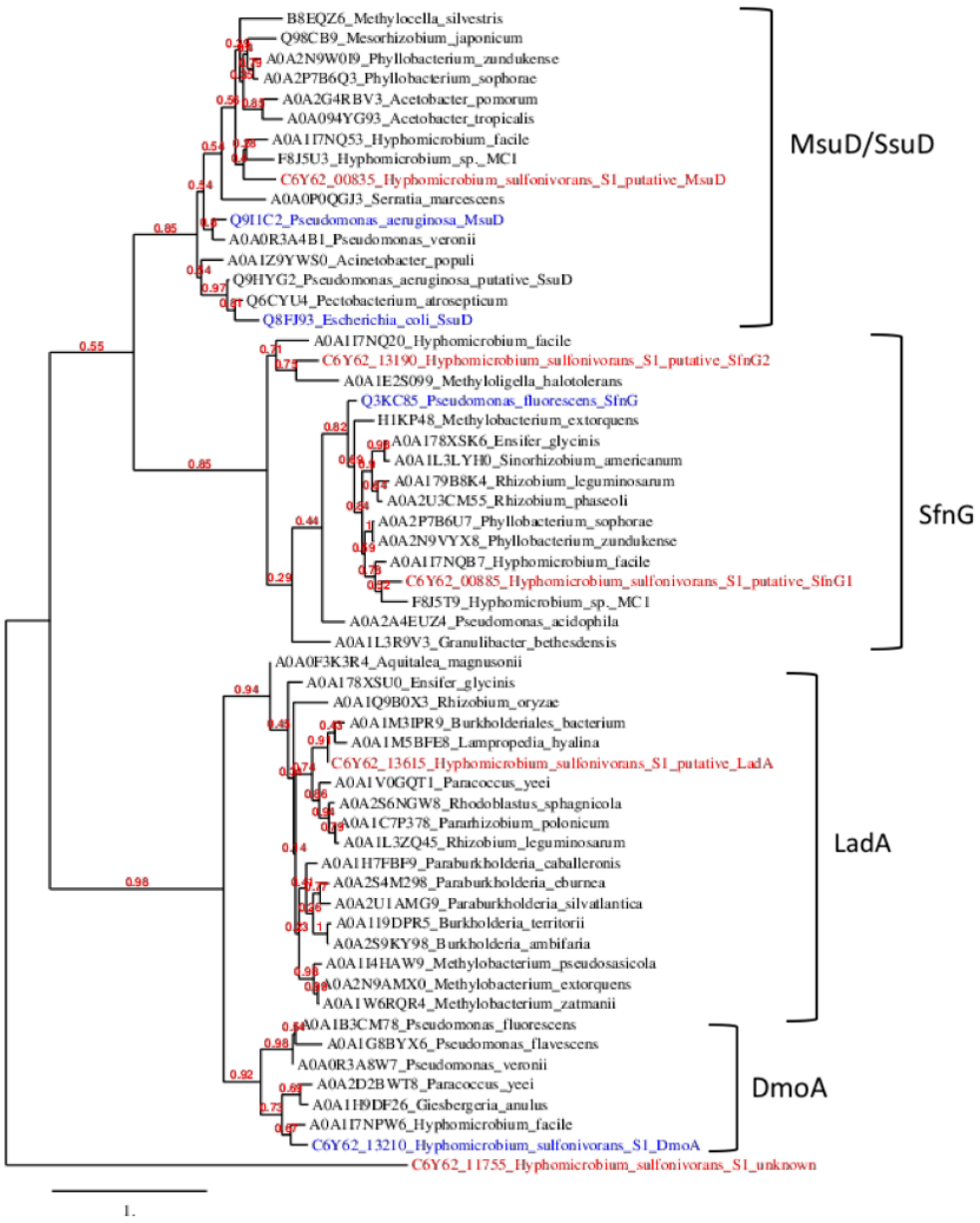


Figure 4.4-1: *Phylogenetic tree of $FMNH_2$-dependent monooxygenase large subunits in* H. sulfonivorans *against those of other bacterial species.* Experimentally characterised reference sequences are labelled BLUE. Uncharacterised sequences from *H. sulfonivorans* are labelled RED. Uncharacterised reference sequences from other bacterial species are labelled BLACK. The confidence of each branch is displayed as a Bootstrap value (RED). UNIPROT Accession numbers are used for bacterial protein sequences except *H. sulfonivorans*, which uses NCBI.

4.4.2 Conclusions

The relatively close clustering between the two SfnG-like sequences (C6Y62_00885 and C6Y62_13190) and the experimentally characterised SfnG from *P. fluorescens* (Wicht, 2006) is a good indicator that they are closely related enzymes. It also supports the earlier assignment of both of these enzymes as $DMSO_2$ monooxygenases made in Chapter 3 based on homology and KEGG orthology.

It may be encouraging to note that the putative MsuD from *H. sulfonivorans* appears to be more similar to the MsuD-type MSA monooxygenase of *P. putida* than the SsuD and putative SsuD-type alkanesulfonate monooxygenases of *E. coli* and *P. aeruginosa*, supporting the earlier assignment of this enzyme as an MSA monooxygenase.

LacA appears to be more closely related to DmoA than any of the other *H. sulfonivorans* monooxygenases, though differentiation between the DmoA and LacA-like sequences suggests that LacA has a distinct metabolic function. It is unfortunate that, as yet, no LadA-like proteins have been experimentally characterised as this makes it difficult to predict the role of the enzyme in the assimilation of $DMSO_2$ as a sulfur source, if indeed it plays any role at all.

In general, the phylogenetics of these $FMNH_2$-dependent monooxygenases does not conflict with the predictions made in Chapter 3 that these proteins are likely to act as the large subunits of a DMS monooxygenase (DmoA), $DMSO_2$ monooxygenase (SfnG1 and SfnG2) and MSA monooxygenase (MsuD) but brings us no closer to understanding the purpose of LadA.

4.5 Discussion

The goals of this Chapter have been to use functional genomics to identify the enzymes of the alternate sulfur-specific pathway of $DSMO_2$ utilisation, identify the uncharacterised enzymes of methylotrophic $DMSO_2$ metabolism in *H. sulfonivorans*, improve our understanding of the downstream metabolism of $DMSO_2$ and finally to identify other genes or systems of interest for future study.

This has led to the identification of three putative MSC monooxygenases consistent with the $DMSO_2$ oxidation pathway put forward in Chapter 1, including two putative SfnFG-type $DMSO_2$ monooxygenases and a putative MsuDE-type MSA monooxygenase. Together with the DmoAB-type DMS monooxygenase characterised by Boden *et al.* (2011), these have been divided into two distinct metabolic pathways discussed below.

4.5.1 $DMSO_2$ metabolism as a sole sulfur source

At the beginning of the project it was thought that the DMS oxidation pathway represented the sole pathway of MSC metabolism in *H. sulfonivorans*. However, the phenotyping performed in Chapter 3 revealed that *H. sulfonivorans* was also capable of utilising MSA, DMS, DMSO and $DMSO_2$ as a sole source of sulfur via a DMS monooxygenase-independent mechanism.

The functional genomics performed in this chapter found that two subunits of an SfnFG-type $DMSO_2$ monooxygenase (SfnG1 and SfnF1), and the large subunit of an MSA monooxygenase (MsuD), are massively upregulated in the proteome of *H. sulfonivorans* when cultivated on $DMSO_2$ as a sole sulfur source versus sulfate, but that all three of these proteins appear to be repressed when *H. sulfonivorans* was cultivated on $DMSO_2$ as a either a sole carbon source or sole carbon and sulfur source. Furthermore, all three of these proteins appear to be encoded by the same gene cluster; C6Y62_00835:00920, that also includes a putative alkanesulfonate ABC transporter which is similarly induced under the same conditions.

The phylogenetic analysis performed in Section 4.4 shows that the large subunits of these monooxygenases are both closely related to their respective, experimentally characterised homologues from other species, suggesting that their tentative assignment as MSC monooxygenases in Chapter 3 may be accurate. On this basis, an alternative pathway of $DMSO_2$ oxidation has been proposed in Figure 4.5-1 that involved the oxidation of $DMSO_2$ to MSIA via SfnFG1, leading to MSA production, followed by the MsuDE-mediated oxidation

of MSA to sulfite. As the phenotyping of *H. sulfonivorans* strains in Chapter 3 demonstrated that the organism can also utilise DMS, DMSO and MSA as a sulfur source in the absence of a functional DmoAB, two more potential enzymes have been included on this pathway – a DMS dehydrogenase and DMSO dehydrogenase – which may double as enzymes of the methylotrophic, reductive metabolism of $DMSO_2$ to DMS for use as a carbon source via the DMS oxidation pathway.

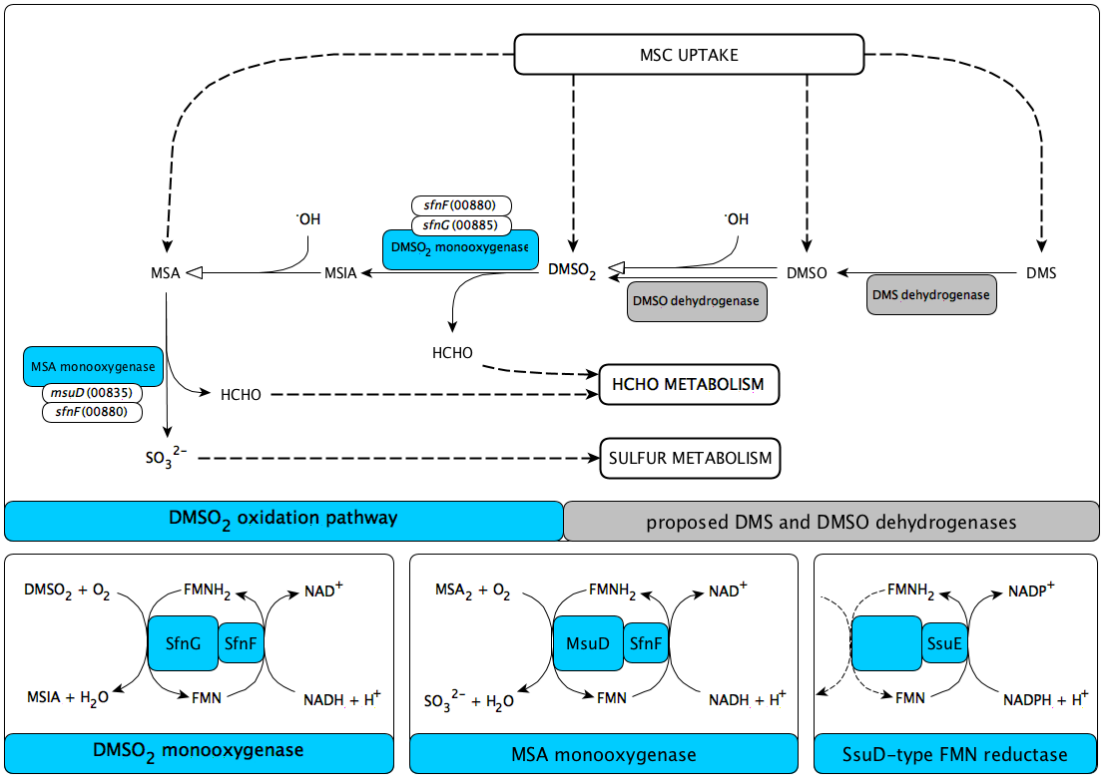


Figure 4.5-1: *MSC metabolism as a sulfur source via the MsuD/SfnG $DMSO_2$ oxidation pathway.* (TOP) Proposed pathway of DMS, DMSO, $DMSO_2$ and MSA metabolism in *H. sulfonivorans* when utilising MSCs as a sole sulfur source, producing formaldehyde (HCHO) for carbon assimilation and sulfite (SO_3^{2-}) for sulfur assimilation. Enzymes of the $DMSO_2$ oxidation pathway are labelled in BLUE, while the unknown enzymes of the prospective DMS reduction pathway are labelled in GREY. Spontaneous reactions between a substrate and hydroxyl residue are labelled with a WHITE arrowhead. (BOTTOM) Proposed reaction mechanisms mediated by the NADH-dependent SfnFG-type $DMSO_2$ monooxygenase (LEFT), NADH-dependent MsuD-SfnF-type MSA monooxygenase (MIDDLE) and an alternative NADPH-dependent mechanism for these enzymes utilising an SsuD-type FMN reductase instead of SfnF (RIGHT).

Note that as only one FMN reductase (SfnF1) has been identified in the C6Y62_00835:00920 gene cluster, it is possible that MsuDE and SfnFG1 share the same small subunit. As the NCBI annotation of SfnF1 in Chapter 3 describes the protein as an NADPH-dependent SsuE-type FMN reductase rather than an NADH-dependent MsuE or SfnF, an alternative reaction scheme for this enzyme has been provided in Figure 4.5-1.

The regulation, role and potential problems of this MsuDE/SfnFG pathway will be discussed in more detail in Chapter 6, but at this stage it represents a strong candidate for the alternate, non-methylotrophic pathway of $DMSO_2$ that was proposed to mediate the assimilatory sulfur metabolism of MSA, $DMSO_2$, DMSO and DMS in the *ΔdmoA* strain.

4.5.2 Methylotrophic $DMSO_2$ metabolism

Despite the identification of a putative $DMSO_2$ oxidation pathway associated with assimilatory sulfur metabolism, the DMS oxidation pathway is still thought to be the key mediator of methylotrophic $DMSO_2$ in *H. sulfonivorans*. The phenotyping of *H. sulfonivorans* described in Chapter 3 demonstrated that the organism is capable of utilising $DMSO_2$ as a sole carbon source using a mechanism that is dependent on the activity of a DmoAB-type DMS monooxygenase.

A search for the genes encoding this enzyme in the transcriptomics data has found that both the *dmoA* and *dmoB* genes are massively upregulated compared to a non-MSC control condition when *H. sulfonivorans* grows on $DMSO_2$, while the proteomics data suggests that this induction of the DmoA and DmoB subunits is far more pronounced in the absence of an alternative methylotrophic carbon source; i.e. MeOH. However, the identity of the three other proposed enzymes of methylotrophic $DMSO_2$ metabolism – $DMSO_2$ reductase, DMS reductase and MT oxidase – remains unknown.

In an unexpected development, it was found that the putative $DMSO_2$ monooxygenase SfnFG2 encoded ~3 kb upstream of the *dmoA* gene is also induced in the *H. sulfonivorans* proteome and transcriptomes when cultivated in the presence of $DMSO_2$, and that this induction is also far greater in the absence of an alternative methylotrophic carbon source; i.e. MeOH. A closer inspection of the omics data revealed that this induction extends to several other genes up and downstream of *dmoA,* which will be referred on here as the *dmoA*/*sfnFG2* gene cluster (C6Y62_13175:13255), suggesting the cluster as a whole may play an important role in methylotrophic $DMSO_2$ metabolism. The arrangement of this gene cluster is displayed in Figure 4.3-9.

A functional analysis of the cluster in Section 4.3 using NCBI annotation and KEGG orthology suggests that the cluster encodes a putative SoxCD sulfur oxidation system enzyme, several *rib* genes of cofactor biosynthesis, a TetR family transcriptional regulator and several transporters from the ABC, porin and MFS families. The specific functions of these transporters are unknown, but as another ABC transporter has already been implicated in the import of extracellular MSCs into *E. coli* for assimilatory sulfur metabolism (Eichorn *et al.*, 2000) it is not unreasonable to propose that one or more of the *dmoA*/*sfnFG2* cluster transporters could mediate the import of $DMSO_2$ into the cell for methylotrophic $DMSO_2$ assimilation.

Returning to SfnFG2, it was initially thought that this putative $DMSO_2$ monooxygenase could be another enzyme of the sulfur assimilatory $DMSO_2$ pathway that facilities the utilisation of MSCs as a sulfur source in the absence of a functional DmoAB. However, the relatively poor induction of SfnFG2 seen in the proteome of *H. sulfonivorans* cultivated on a sole sulfur source of $DMSO_2$, combined with the close proximity of the *sfnG2* and *sfnF2* genes to *dmoA* and *dmoB*, suggests that the enzyme is unlikely to play a dominant role in the assimilation of $DMSO_2$ as a sulfur source and is more likely to be involved in methylotrophic $DMSO_2$ metabolism instead.

The co-induction of a $DMSO_2$ monooxygenase with a DMS monooxygenase is counterintuitive given that DMS oxidation and $DMSO_2$ oxidation are thought to be mutually exclusive processes; the DMS oxidation pathway is thought to produce hydrogen sulfide via MT oxidation while the $DMSO_2$ oxidation pathway is thought to yield sulfite via MSA oxidation.

Assuming that SfnFG2 does indeed play a role in methylotrophic $DMSO_2$ metabolism this presents at least three possibilities: (i) that a second $DMSO_2$ oxidation pathway involving SfnFG2 is expressed in parallel with the DmoAB DMS oxidation pathway, (ii) that SfnFG2-mediated $DMSO_2$ oxidation only leads to partial $DMSO_2$ dissimilation, or (iii) that the proposed function of SfnFG2 as a $DMSO_2$ monooxygenase is incorrect and the enzyme plays some other role. A visual representation of these hypotheses is presented in Figure 4.5-2 and a critique of each hypothesis is provided below:

Hypothesis 1: The $DMSO_2$ oxidation pathway works in parallel with DMS oxidation

The first hypothesis posits that the $DMSO_2$ monooxygenase SfnFG2 oxidises $DMSO_2$ to MSIA as part of a $DMSO_2$ oxidation pathway that acts in parallel with reductive $DMSO_2$

metabolism which generates DMS for methylotrophic DMS metabolism via the DMS monooxygenase. The presumption is that another as yet unidentified MSA monooxygenase is also present in *H. sulfonivorans* in addition to the MsuDE-type monooxygenase of the MsuDE/SfnFG1 $DMSO_2$ oxidation pathway. This SfnFG2-mediated $DSMO_2$ oxidation pathway would be expected to yield two molecules of formaldehyde and one molecule of sulfite for each $DMSO_2$ molecule that enters the pathway.

However, the generation of the C1 compound formaldehyde is inconsistent with the inability of the Δ*dmoA* strain to utilise $DMSO_2$ as a sole carbon source. It follows that if more than one pathway of methylotrophic $DMSO_2$ metabolism is induced when *H. sulfonivorans* was cultivated on $DMSO_2$ as a sole carbon source then the *dmoA* gene disruption mutant would still be capable of metabolising DMSO as a carbon source. An exception to this would be if the specific mechanism of MSA utilisation prevents the downstream metabolism of its products, as discussed above for the MsuD/SfnG1 $DMSO_2$ oxidation pathway. This may include the autoinhibition of the enzyme, inefficiency of the pathway or the overproduction of toxic products exceeding the organism's capacity to remove them.

Hypothesis 2: $DMSO_2$ oxidation leads to MSA excretion

The second hypothesis is that the $DMSO_2$ monooxygenase SfnFG2 oxidises $DMSO_2$ to MSIA, but that there is no downstream mechanism for the catabolism of either MSIA or MSA to produce formaldehyde for methylotrophic carbon assimilation. This requires that the carbon of any MISA or MSA produced is ultimately excreted from the cell rather than being used as a methylotrophic substrate for carbon assimilation.

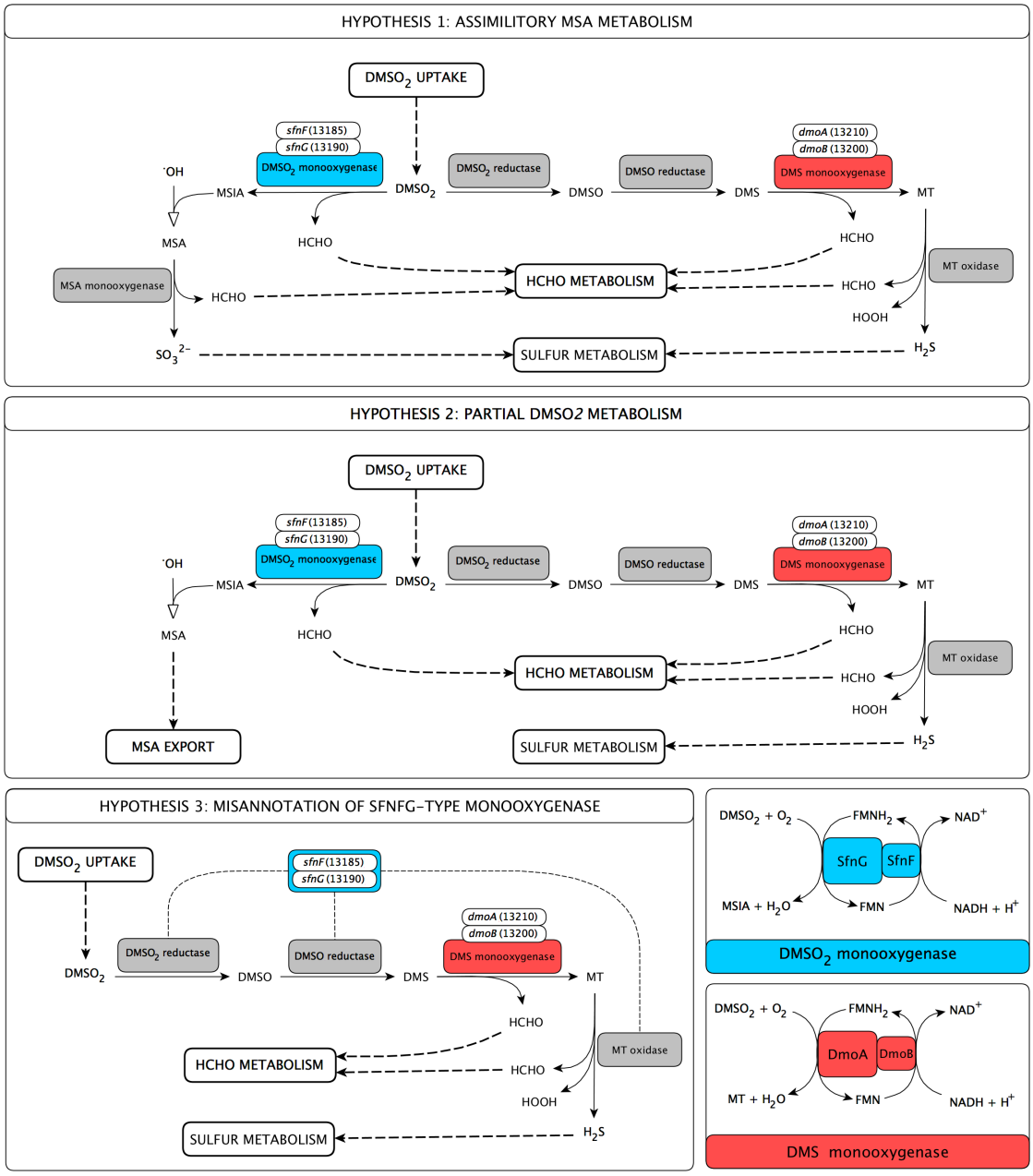


Figure 4.5-2: *Potential mechanisms of $DMSO_2$ metabolism in* H. sulfonivorans *as a carbon source via the DmoAB-type DMS monooxygenase and the putative SfnFG-type $FMNH_2$-depependent monooxygenase.* (TOP) Option 1, in which the SfnFG-type monooxygenase mediates an assimilatory $DMSO_2$ oxidation pathway to generate formaldehyde and sulfite. (MIDDLE) Option 2, in which the SfnFG-type monooxygenase mediates a partial assimilatory/dissimilatory $DMSO_2$ oxidation pathway to generate formaldehyde and MSA. (BOTTOM LEFT) Option 3, which the SfnFG-type monooxygenase has been misannotated and may instead mediate $DMSO_2$ reduction, DMSO reduction and/or MT oxidation. (BOTTOM RIGHT) Proposed reaction mechanisms of $DMSO_2$ oxidation by the putative SfnFG-type $DMSO_2$ monooxygenase (BLUE) and the DmoAB-type DMS monooxygenase (RED).

Not only is this mechanism quite wasteful, discarding half of the potential carbon present in each molecule of $DMSO_2$ metabolised by SfnFG2, but such a mechanism would still be capable of generating formaldehyde for methylotrophic carbon assimilation. Again, this conflicts with the finding in Chapter 3 that methylotrophic $DMSO_2$ metabolism occurs via the activity of the DmoAB-type DMS monooxygenase unless an extenuating circumstance prevents the formaldehyde produced by SfnFG2 from being used as a sole carbon source.

Hypothesis 3: SfnFG2 mediates reductive $DMSO_2$ metabolism or the DMS oxidation pathway

A third explanation is that the annotation of SfnFG2 as a putative $DMSO_2$ monooxygenase is incorrect and the enzyme actually has a quite different function. The phylogenetics described in Section 4.4 showed that, at minimum, the SfnG2 subunit is a close relative of $FMNH_2$-dependent MSC monooxygenases. It is therefore quite likely that even if SfnFG2 is not a $DMSO_2$ monooxygenase it may still play a role in the oxidation or reduction of other MSCs, especially as the enzyme is heavily induced in response to the methylotrophic growth of *H. sulfonivorans* on $DMSO_2$.

As the position of DMS monooxygenase is already taken by DmoAB, this leaves three of the enzymes of reductive $DMSO_2$ metabolism and the DMS oxidation pathway that have been previously proposed by Borodina *et al.* (2000) and Boden *et al.* (2011): $DMSO_2$ reductase, DMSO reductase and MT oxidase. All of this is consistent with the previous finding that methylotrophic $DMSO_2$ metabolism occurs via a DmoAB-type DMS monooxygenase.

Enzyme assays performed by Borodina *et al.* (2000) have previously found that *H. sulfonivorans* S1 cell lysate exhibits NADH-dependent $DMSO_2$ and DMSO reductase activity, found to be much greater in the lysate's membrane fraction than the soluble fraction. Oxygen electrode experiments of *H. sulfonivorans* lysate was also found to mediate the NADH-independent oxidation of MT oxidase consistent with the presence of a MT reductase. This suggests that the *H. sulfonivorans* DMSO and $DMSO_2$ reductases are both NADH-dependent and membrane bound, while the MT oxidise is NADH-independent and soluble.

While SfnFG2 is indeed predicted to be NADH-dependent, $FMNH_2$-dependent monooxygenases are soluble enzymes and no distinctive binding domain has been discovered in either the SfnF2 or SfnG2 amino acid sequences. If the enzyme assays performed by Borodina *et al.* (2000) are correct, then it may be quite unlikely that SfnFG is either the $DMSO_2$ reductase, DMSO reductase or MT oxidise. This does not preclude the involvement in SfnFG2 in either DMSO reduction, $DMSO_2$ reduction or MT oxidation, but it does indicate that either

SfnFG2 uses a distinct mechanism is functionally distinct from the other $FMNH_2$-dependent monooxygenases or our limited understanding of the rest of the reductive $DMSO_2$ pathway enzymes is incorrect.

DMS monooxygenase pathway is essential for methylotrophic $DMSO_2$ metabolism

Although the precise function of SfnFG2 in the methylotrophic metabolism of $DMSO_2$ remains unknown, each of the three scenarios presented above is largely consistent with the previous hypothesis made at the end of Chapter 3, that "The methylotrophic DMS monooxygenase pathway is the sole mechanism by which *H. sulfonivorans* utilises DMSO2 as a sole carbon source".

However, the discovery that the $DMSO_2$ oxidation pathways is also induced in conjunction with the DMS monooxygenase pathway when *H. sulfonivorans* was cultivated on $DMSO_2$ as a sole carbon and sulfur source suggests that the hypothesis may not be entirely accurate; formaldehyde produced by the $DMSO_2$ monooxygenase pathway is likely to also be assimilated as a s sole carbon source. The hypothesis has therefore been updated to resolve this conflict:

The methylotrophic DMS monooxygenase pathway is an essential mechanism for the utilisation of $DMSO_2$ as a sole carbon source

Potential methods for testing this hypothesis, establishing the function of SfnFG2, and determining the identities of outstanding members of the DMS monooxygenase pathway will be discussed in Chapter 6.

Chapter 5:

Comparative genomics of MSC metabolism in *Hyphomicrobium species*

5: Comparative genomics of MSC metabolism in *Hyphomicrobium* species

5.1 *Hyphomicrobium* species, model organisms of MSC metabolism

In addition to *H. sulfonivorans* S1, this project has also examined MSC metabolism in four other *Hyphomicrobium* species which have either draft or complete genomes available for study: *H. denitrificans* ATCC 51888, *H. facile* Bras 3, *H. methylovorum* Bras1 and *H.* sp. VS. What makes these organisms particularly interesting is that they have all previously been shown to utilise other MSCs as a sole carbon source, either instead of or in addition to $DMSO_2$. It is therefore possible that these organisms make use of different molecular mechanisms than those found in *H. sulfonivorans* to degrade these substrates.

The intention of the work presented in this chapter is to use functional genomics to compare and contrast the metabolism of MSCs in *H. sulfonivorans*, with that of other *Hyphomicrobium* species. The potential benefits of this strategy are twofold: Firstly, to gain a better understanding of how MSC metabolism varies between different *Hyphomicrobium* species, and secondly, to use this information to refine our existing models of MSC metabolism in *H. sulfonivorans*.

An overview of the research performed in each model *Hyphomicrobium* species is described below, to put the functional genomics of these species into the context of their existing literature.

5.1.1 Overview of model *Hyphomicrobium* species

Hyphomicrobium denitrificans *ATCC 51888*

Hyphomicrobium denitrificans strain ATCC 51888 (TK 0415) has been studied as a model organism of methylotrophic metabolism for half a century. The organism was first described by Harder *et al.* (1973) as "*Hyphomicrobium* X" and would be studied as a model methylotroph for two decades before being identified as a serine cycle methylotroph by Ukurami *et al.* (1995), when it was assigned the novel species name *H. denitrificans*.

The complete genome of *H. denitrificans* ATCC 51888 has already been sequenced by Brown *et al.* (2011) as part of a study into the prosthecate forming Alphaproteobacteria, making the organism an ideal model for comparative and functional genomics. This appears to be an

opinion shared by other researcher groups, as *H. denitrificans* has recently been studied as a model organism of organosulfur compound metabolism by Koch and Dahl (2018).

The work by Koch and Dahl, briefly discussed in earlier chapters, involved a combination of mutagenesis and functional genomics to identify a novel sulfur oxidation system, induced in response to the growth of the organism on DMS as a sole carbon source. This heterodisulfide reductase (Hdr) system is thought to oxidise thiosulfate to sulfite in a process which is critical for the methylotrophic growth of *H. denitrificans* on DMS. Fortunately, the research performed by Koch and Dahl (2018) has focused on inorganic sulfur metabolism rather than MSC oxidation, so despite some overlap it does not appear to have superseded the work described in this chapter.

Hyphomicrobium species *VS*

Hyphomicrobium sp. VS is a species that was isolated from activated sewage sludge by enrichment in the presence of DMS (Pol *et al.*, 1994). Pol *et al.* (1994) attempted to cultivate the species on a variety of methylotrophic substrates and found that the strain was capable of utilising DMS and DMSO as a carbon source, as well as several other methylotrophic compounds including MeOH, formaldehyde, formate and methylamines. Oxygen electrode assays suggested that cells of *H.* sp. VS were able to respire MT and DMS when cultivated on DMS, or MT, DMS and DMSO when cultivated on DMSO.

Several enzyme assays were also performed by Pol *et al.* (1994) on cell-free extracts derived from *H.* sp. VS biomass which had been cultivated on DMS. These were able to identify the presence of MT oxidase activity but failed to find any activity for the as yet uncharacterised DMS monooxygenase believed to be involved in DMS oxidation. This MT oxidase would later be characterised by Eyice *et al.* (2017) as an MtoX-type NADH-independent MT oxidase that oxidises MT to produce formaldehyde, hydrogen peroxide and hydrogen sulfide.

To briefly summarise the characterisation of this MT oxidase, Ozge *et al.* (2018) began by purifying the enzyme from *H.* sp. VS biomass cultivated on DMS. Enzyme assays of the purified protein were then performed using a combination of oxygen electrode experiments, gas chromatography and biochemical assays of formaldehyde formation, demonstrating the enzyme was capable of oxidising MT to generate formaldehyde. The amino acid sequence of this MtoX-type MT oxidase was then identified by peptide sequencing the enzyme's N-terminal domain and examining internal peptides from the enzyme using mass spectrometry.

These peptide sequences were then used to identify the nucleotide sequence of the *mtoX* gene in the *H.* sp. VS genome, which had been sequenced by Ozge *et al.* (2018) by Illumina genome sequencing.

Hyphomicrobium facile Bras3 *and Hyphomicrobium methylovorum* Bras1

The final two *Hyphomicrobium* species are *Hyphomicrobium methylovorum* strain Bras1 and *Hyphomicrobium facile* strain Bras3, both of which have been isolated from the rhizosphere of *Brassica oleracea* (Warwickshire, UK) (Eyice and Schäfer, 2015). *Hyphomicrobium methylovorum* Bras1 was isolated from *Brassica oleracea* by performing an enrichment experiment with a sole carbon source of MeOH, while *Hyphomicrobium facile* Bras3 was enriched with a sole carbon source of DMS. Growth assays performed by Eyice and Schäfer (2015) suggested that both *H. methylovorum* Bras1 and *H. facile* Bras3 are capable of utilising either MeOH, DMS or DMSO as a sole carbon source. The genomes of both strains would later be sequenced by Illumina sequencing and assembled into draft genomes (Schäfer *et al.*, unpublished). The phenotypes and available genomes of both these organisms make them ideal model organisms for the study of MSC metabolism in *Hyphomicrobium* species.

5.1.2 Preliminary genotyping of *Hyphomicrobium* species

The genotyping of *H. sulfonivorans* in Chapter 3 demonstrated that the organism lacks the MtoX-type MT oxidase previously characterised in *H.* sp. VS (Eyrice *et al.*, 2018), raising the possibility that different *Hyphomicrobium* species make use of different mechanisms to metabolise MSCs. It may therefore be interesting to search the genomes of several *Hyphomicrobium* species for known enzymes of MSC metabolism identified in other bacterial species to make a basic assessment of which metabolic pathways may be present in each strain.

DMS oxidation pathway

The two enzymes of MSC metabolism that have been identified in *Hyphomicrobium* species are the DMS monooxygenase of *H. sulfonivorans* (Boden *et al.*, 2011) and the MT oxidase of *H.* sp. VS (Eyice *et al.*, 2017). However, the *mtoX* gene encoding MT oxidase in *H.* VS appears to be absent from the *H. sulfonivorans* genome (see Section 3.4). The first step in examining MSC metabolism in these other *Hyphomicrobium* species was to genotype them for the presence or absence of these enzymes. BLAST searches of the *H. denitrificans* ATCC

51888, *H. facile* Bras3, H. *methylovorum* Bras1 and *H.* VS genomes were performed using DmoA (NCBI Accession ADU77278.1) and MtoX (NCBI Accession ATJ26742.1) as query sequences (see Table 5.2-1). Based on the presence of homologues for DMS monooxygenase and MT oxidase, these *Hyphomicrobium* species can be divided into three genotypes. The three species *H. facile*, *H. methylovorum* and *H.* sp. VS appear to have a strong homologue of both enzymes present in their genome. In contrast, *H. sulfonivorans* only has the DMS monooxygenase and *H. denitrificans* only has the MT oxidase.

Table 5.1-1: *BLAST search for the DMS oxidation pathway in* Hyphomicrobium *species*

	DMS monooxygenase (DmoA) NCBI/ ADU77278.1				MT oxidase (MtoX) NCBI/ ATJ26742.1			
	Present	Highest Identity	Score (Bits)	E value	Present	Highest Identity	Score (Bits)	E value
H. denitrificans ATCC 51888	NO	*25%*	*28.9*	*0.71*	**YES**	***97%***	***667***	***0.0***
H. facile Bras3	**YES**	***86%***	***849***	***0.0***	**YES**	***89%***	***794***	***0.0***
H. methylovorum Bras1	**YES**	***81%***	***805***	***0.0***	**YES**	***99%***	***870***	***0.0***
H. sulfonivorans S1	**YES**	***~100%***	***970***	***0.0***	NO	*28%*	*31.2*	*0.13*
H. species VS	**YES**	***81%***	***805***	***0.0***	**YES**	***99%***	***899***	***0.0***

$DMSO_2$ oxidation pathway

BLAST searches for members of the proposed $DMSO_2$ oxidation pathway in *H. sulfonivorans* S1 were performed for each of the *Hyphomicrobium* genomes, using the same query sequences as described in the genome analysis of *H. sulfonivorans* performed in Chapter 3. This included the SfnFG-type $DMSO_2$ monooxygenase SfnG subunit from *Pseudomonas fluorescens* (Wicht, 2016) and the MsuDE-type MSA monooxygenase from *Pseudomonas aeruginosa* (Kertesz *et al.*, 1999). This has identified an SfnFG-like monooxygenase in the draft genomes of *H. facile*, *H. methylovorum* and *H.* sp. VS (see Table 5.1-2), each being a distinct sequence to the DmoA homologues described above.

However, a search for an MsuD-type monooxygenase failed to identify a strong homologue for any of the four *Hyphomicrobium* species (see Table 5.1-2). In response, further BLAST searches were performed for the alpha subunit of another bacterial MSA

monooxygenase, an MsmABCD-type MSA monooxygenase from *Methylosulfonomonas methylovora* (UNIPROT accession Q9X404) (Marco *et al.* 1999). While the enzyme was absent in *H. denitrificans*, which also has no MsuDE-type monooxygenase, strong homologues were found in the draft genomes of *H. facile*, *H. methylovorum* and *H.* sp. VS (see Table 5.1-3).

The MsmABCD-type transporter of *Methylosulfonomonas methylovora* has previously been found alongside an MsmEFGH-type ABC transporter, thought to facilitate the transport of extracellular MSA into the cell (Jamshad *et al.*, 2006). A search for the MsmE subunit of this transporter (UNIPROT accession Q9X402) in the *Hyphomicrobium* genomes was also performed but was unable to identify any significant MsmE candidates (see Table 5.1-3), suggesting that this transporter is absent in these *Hyphomicrobium* species.

Table 5.1-2: *BLAST search for $DMSO_2$ oxidation pathway enzymes in* Hyphomicrobium *species*

	$DMSO_2$ monooxygenase large subunit (*Pseudomonas fluorescens* SfnG) - UNIPROT/ Q3KC85				MSA monooxygenase large subunit (*Pseudomonas aeruginosa* MsuD) - UNIPROT/ Q9I1C2			
	Present	Highest Identity	Score (Bits)	E value	Present	Highest Identity	Score (Bits)	E value
H. denitrificans ATCC 51888	NO	*no significant hits*			NO	*30%*	*40*	*1e-04*
H. facile Bras3	**YES**	***59%***	***349***	***e-138***	NO	*28%*	*122*	*2e-32*
H. methylovorum Bras1	**YES**	***60%***	***452***	***e-160***	NO	*31%*	*117*	*5e-31*
H. sulfonivorans S1	**YES**	***73%***	***554***	***0.0***	**YES**	***67%***	***488***	***1e-172***
H. species VS	**YES**	***61%***	***452***	***1e-160***	NO	*31%*	*133*	*2e-36*

Table 5.1-3: *BLAST search for Msm-type MSA monooxygenase and ABC transporter in* Hyphomicrobium *species*

	MSA monooxygenase subunit A (*Methylosulfonomonas methylovora*, MsmA) UNIPROT/ Q9X404				ABC MSA transporter subunit E (*Methylosulfonomonas methylovora*, MsmE) UNIPROT/ Q9X402			
	Present	Highest Identity	Score (Bits)	E value	Present	Highest Identity	Score (Bits)	E value
H. denitrificans ATCC 51888	NO	*no significant hits*			NO	*no significant hits*		
H. facile Bras3	**YES**	***83%***	***730***	***0.0***	NO	*no significant hits*		
H. methylovorum Bras1	**YES**	***85%***	***757***	***0.0***	NO	*no significant hits*		
H. sulfonivorans S1	NO	*no significant hits*			NO	*no significant hits*		
H. species VS	**YES**	***84%***	***759***	***0.0***	NO	*no significant hits*		

Conclusions

The genotyping of MSC metabolism in these *Hyphomicrobium* species points towards three distinct genotypes. The first genotype applies to *H. denitrificans* ATCC 51888, which contains an MtoX-type MT oxidase but lacks a DmoAB-type DMS monooxygenase, SfnFG-type $DMSO_2$ monooxygenase, MsuDE-type MSA monooxygenase and MsmABCD-type MSA monooxygenase. The second genotype applies to *H. facile* Bras3, *H. methylovorum* Bras1 and *H.* species VS, which contain a DmoA, MtoX, SfnG and MsmA but lack an MsuD. The third and final genotype applies to *H. sulfonivorans* S1, which contains a DmoA, SfnG and MsuD but lacks an MtoX and MsmA. For clarity, a summary of these genotypes is provided in Table 5.1-4 below.

Table 5.1-4: *Summary of the three genotypes of MSC metabolism exhibited in five* Hyphomicrobium *species*

Species	Gene presence				
	DmoA	MtoX	SfnG	MsuD	MsmA
H. denitrificans ATCC 51888	-	**YES**	-	-	-
H. facile Bras3	**YES**	**YES**	**YES**	-	**YES**
H. methylovorum Bras1	**YES**	**YES**	**YES**	-	**YES**
H. sulfonivorans S1	**YES**	-	**YES**	**YES**	-
H. species VS	**YES**	**YES**	**YES**	-	**YES**

While the functional genomics of MSC metabolism in *H. sulfonivorans* shows that this may be far from an exhaustive list of the enzymes present in that organism, let alone all five *Hyphomicrobium* species, it does provide a good starting point for a comparative analysis between the different members of the *Hyphomicrobium* genus. If these putative enzymes do indeed play an important role then we would expect to see this variation in genotype to also be expressed in the phenotypes of these organisms when cultivated on different MSCs, and may even be able to make some tentative predictions for each organism based on the presence or absence or particular genes. This will be explored by the phenotyping of *Hyphomicrobium* species that is described in Section 5.2.

5.2 Phenotyping *Hyphomicrobium* species

The first experimental task of this chapter has been to confirm the range of MSCs that the four model *Hyphomicrobium* species can utilise as a carbon and/or sulfur source. This has been performed by cultivating *H. denitrificans* ATCC 51888, *H. facile* Bras3, *H. methylovorum* Bras1 and *H.* sp. VS following the same methodology as previously seen for the phenotyping of MSC utilisation in *H. sulfonivorans* S1, described in Section 3.2.

5.2.1 Carbon assimilation phenotyping

H. denitrificans ATCC 51888*, H. facile* Bras3*, H. methylovorum* Bras1 and *H.* sp. VS were cultivated in triplicate on minimal media with a sole carbon source of either $DMSO_2$, DMSO, DMS or MSA against positive condition of MeOH and a negative control condition without carbon. All cultures were supplemented with sulfate as a replete inorganic sulfur source. As DMS toxicity has previously been found to inhibit bacterial growth at concentrations above 1mM, it was decided that the MSCs would be introduced to *H. sulfonivorans* incrementally via cultivation in fed-batch.

Hyphomicrobium cultures were inoculated in parallel for each condition in triplicate, then cultivated for 144 hrs with sampling of the cultures at 0, 48, 96, 120 and 144 hrs for each strain except *H. denitrificans*, which was not sampled at 120 hrs. Growth assays were performed by measuring sample optical density at 540 nm. Cultures were then supplemented with additional carbon source in an equal volume of fresh culture, maintaining a consistent culture volume throughout the experiment. The results of this experiment demonstrate that while each strain showed strong growth on the positive control condition of MeOH, relative to the null sulfur negative control, there is substantial variation in the MSCs that can be utilised as a sole carbon source by the different *Hyphomicrobium* species.

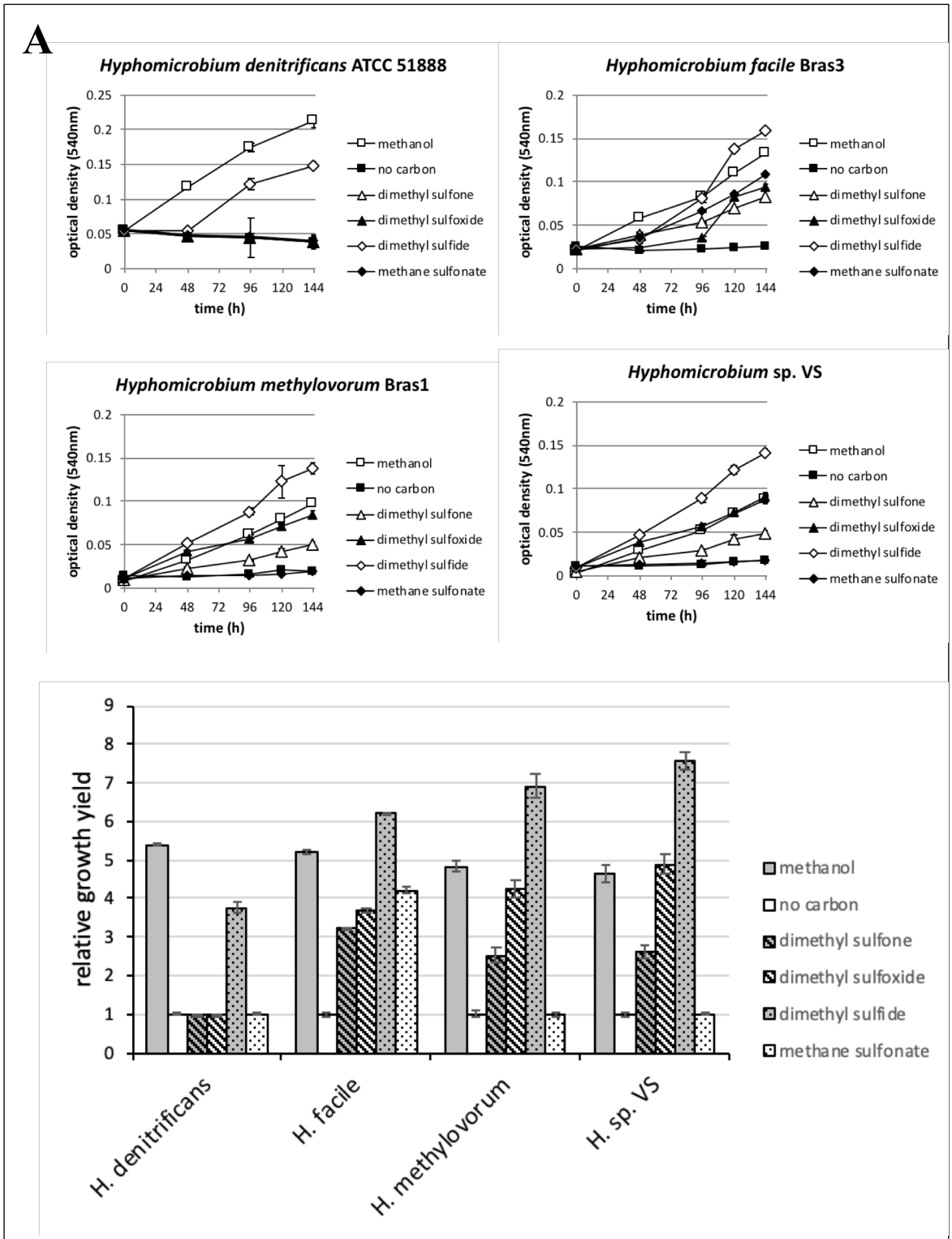


FIGURE 5.2-1: Hyphomicrobium *carbon source phenotyping.* (A) Growth curves of *Hyphomicrobium* species in fed-batch culture on a range of carbon sources, supplemented with additional substrate at each sampling point. (B) Relative growth yield of *Hyphomicrobium* species on various carbon sources at final sampling after 144 h fed-batch cultivation, relative to OD_{540} of the no carbon control condition for each species. All error bars are displayed in standard deviation for triplicate cultures.

H. facile was able utilise all four of the MSCs tested and was the only species capable of utilising MSA as a sole carbon source. *H. methylovorum* and *H.* species VS were both able to use $DMSO_2$, DMSO and DMS, while *H. denitrificans* was only able to utilise DMS and neither $DMSO_2$ or DMSO. A summary of these results has been combined with the previous phenotyping data for the *H. sulfonivorans* S1 WT and displayed in Table 5.2-1.

Table 5.2-1: *Growth profile of* Hyphomicrobium *species on various carbon sources.*

	MeOH	$DMSO_2$	DMSO	DMS	MSA
H. denitrificans ATCC 51888	++	-	-	+	-
H. facile Bras3	++	+	+	++	+
H. methylovorum Bras1	++	+	++	+++	-
H. sp. VS	++	+	++	+++	-
H. sulfonivorans S1	++	+	-	-	-

5.2.2 Sulfur assimilation phenotyping

H. denitrificans ATCC 51888*, H. facile* Bras3*, H. methylovorum* Bras1 and *H.* species VS were cultivated on a range of MSCs as a sole sulfur source to assess the breadth of compounds that each of these *Hyphomicrobium* species are capable of utilising. Each strain was cultivated in triplicate on minimal media with a carbon source of MeOH and a sole sulfur source of either $DMSO_2$, DMSO, DMS, MSA, sodium sulfate or a negative control condition without sulfur. Note that, as was also the case for the phenotyping of *H. sulfonivorans* (see Section 3.2), the carbon:sulfur ratio of the media varied between 100:1 (positive and negative control), 101:1 (MSA) and 102:1 ($DMSO_2$, DMSO and DMS), but that this 1-2% difference has been considered nominal for the purpose of assessing net bacterial growth.

Each strain was inoculated in triplicate cultures for each condition in parallel, then cultivated for 50 h with sampling of the cultures at 0, 12, 36 and 50 h (see Figure 5.2-2). After 50 h growth, each strain had shown clear growth on sulfate when compared to the negative, no sulfur control condition. For the MSCs *H. facile* was able utilise $DMSO_2$, DMS and MSA as a sole sulfur source, *H. methylovorum* and *H.* sp. VS were both able to use $DMSO_2$, DMSO and DMS, while *H. denitrificans* was only able to utilise DMS. A summary of these results has been combined with the previous phenotyping data for the *H. sulfonivorans* S1 WT and displayed in Table 5.2-2.

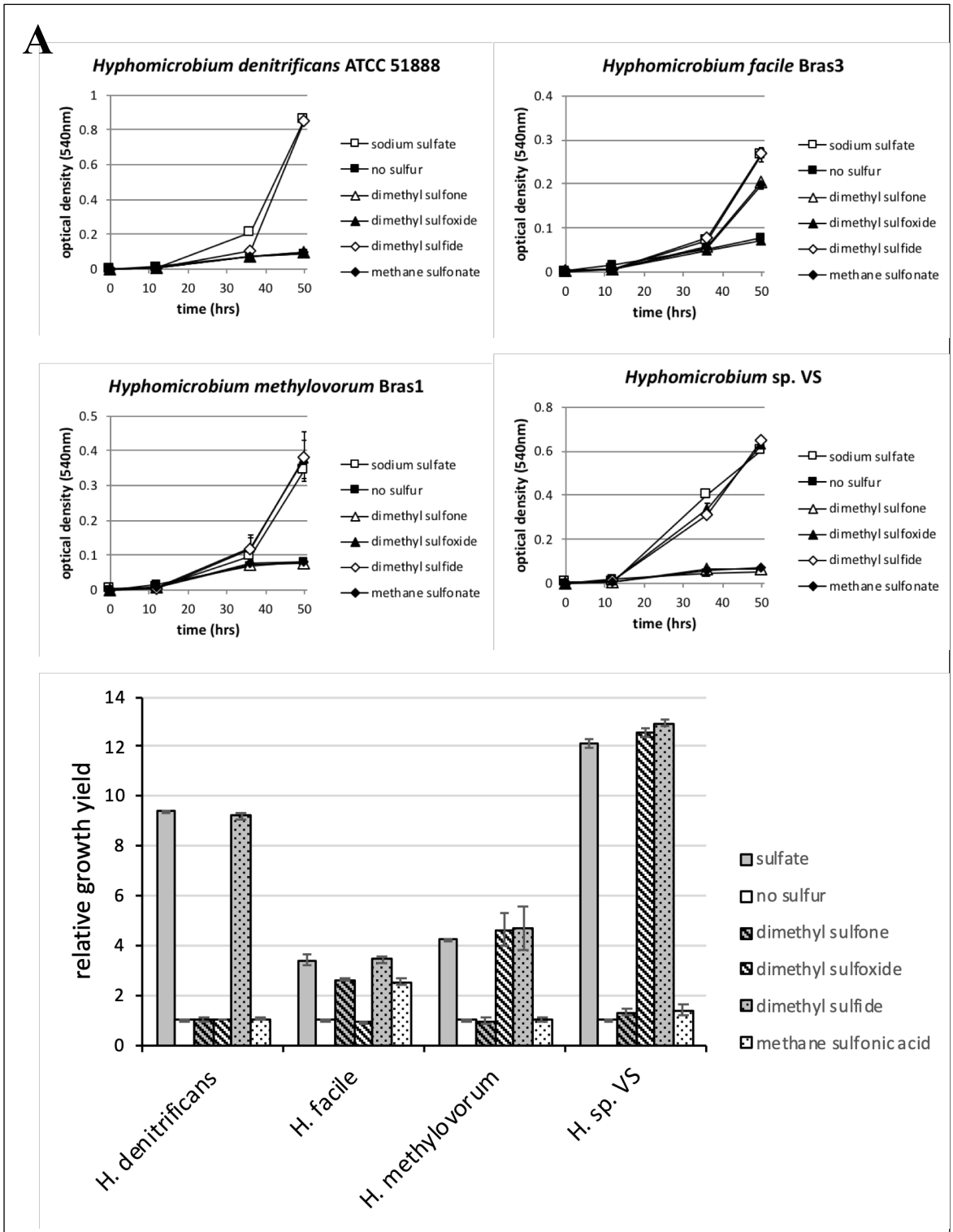


FIGURE 5.2-2: Hyphomicrobium *sulfur source phenotyping.* (A) Growth curves of *Hyphomicrobium* species in fed-batch culture on a range of sulfur sources, supplemented with additional substrate at each sampling point. (B) Relative growth yield of *Hyphomicrobium* species on various sulfur sources at final sampling after 50 h fed-batch cultivation, relative to the no sulfur control condition of each species. All error bars are displayed in standard deviation for triplicate cultures.

Table 5.2-2: *Growth profile of Hyphomicrobium species on various sulfur sources.*

	Na_2SO_4	$DMSO_2$	DMSO	DMS	MSA
H. denitrificans ATCC 51888	++	-	-	++	-
H. facile Bras3	++	++	-	++	++
H. methylovorum Bras1	++	+	++	++	-
H. species VS	++	+	++	++	+/-
H. sulfonivorans S1	++	++	+++	+	++

5.2.3 Conclusions

Each of the five *Hyphomicrobium* species can utilise at least one MSC as both a sole carbon and sole sulfur source, but there is substantial variation in the MSC utilisation profiles shown by each species, summarised in Table 5.2-3. Furthermore, this data shows a general trend between the genotype and phenotype of each species; the presence of *mtoX* correlates with the capacity of the organism to utilise DMS as a sole carbon source, while the *dmoA* correlates with $DMSO_2$ utilisation as a carbon source. This may be an indicator that these species are utilising distinct mechanisms to utilise MSCs, leading these species to exhibit a wide variety of different phenotypes.

Table 5.2-3: *Growth profile of Hyphomicrobium species on various sulfur sources.*

	$DMSO_2$	DMSO	DMS	MSA
H. denitrificans ATCC 51888	-	-	carbon/sulfur	-
H. facile Bras3	carbon/sulfur	carbon	carbon/sulfur	carbon/sulfur
H. methylovorum Bras1	carbon/sulfur	carbon/sulfur	carbon/sulfur	-
H. sulfonivorans S1	carbon/sulfur	sulfur	sulfur	sulfur
H. VS	carbon/sulfur	carbon/sulfur	carbon/sulfur	sulfur

To gain a better grasp of these processes, it may be necessary to perform more detailed genotyping of these species to better explore the relationship between genotype and phenotype, and to serve as a framework for functional genomics. Given the limited time and resources

available for studying these strains, it was decided that only two more *Hyphomicrobium* species would be chosen for further analysis.

The most tangible difference between *Hyphomicrobium* species is the disparity in MSC utilisation between *H. denitrificans* and *H. sulfonivorans.* While *H. sulfonivorans* can utilise $DMSO_2$ as a sole carbon source but not DMS, *H. denitrificans* can utilise DMS as a sole carbon source but not $DMSO_2$. This is in stark contrast to the other *Hyphomicrobium* species which can utilise both $DMSO_2$ and DMS as a sole carbon source, so it was a natural choice to study *H. denitrificans*, *H. sulfonivorans* and one of the other three *Hyphomicrobium* species.

The third species chosen for further study was *H. methylovorum*, selected due to lack of previous characterisation, capacity to utilise $DMSO_2$, DMSO and DMS as both sole carbon source and sole sulfur source, and because it contains a copy of both *dmoA* and *mtoX* in its genome.

5.3 Genome sequencing of *Hyphomicrobium methylovorum* Bras1

Performing an omics analysis of MSC metabolism in *H. methylovorum* required an accurate genome to use as a reference for genomics, proteomics and transcriptomics. Previous sequencing of *H. methylovorum* strain Bras1 by the Schäfer lab group (unpublished) generated a data set of Illumina 70bp length paired-end reads and produced a 3,439,435 bp assembly over 7 contigs, with 3,508 predicted features and with any contigs <500bp discarded. However, with the resequencing of the *H. sulfonivorans* S1 genome already underway it was decided that the draft genome of *H. methylovorum* Bras1 would also be resequenced and reassembled to create a higher quality reference genome for genomic and proteomic analyses.

H. methylovorum Bras1 biomass was cultivated and submitted to the MicrobesNG sequencing service for paired-end Illumina sequencing, performed on an Illumina MiSeq. This generated 250 bp paired-end raw reads that were then assembled into a new genome using the SPADES assembly platform (Bankevich *et al.*, 2012), in a hybrid assembly that combined the new 250 bp read data with the existing 70 bp read data collected by the Schäfer lab group (unpublished).

This new genome for *H. methylovorum* S1 had a total of 3,632,673 base pairs split between 398 contigs, with ~94.8% of base pairs located on contigs greater than or equal to 1 kb in length (see Figure 4.3-1). Contigs shorter than 1 kb were considered too small for multi-omics analysis and discarded from the genome, leaving 3,443,230 base pairs spread over 4 contigs. Similar to the *H. sulfonivorans* genome discussed earlier, this means a loss of ~5.2% from the assembly and it should be kept in mind that the draft is likely to represent an incomplete genome. However, this was deemed a necessary compromise to prevent the misannotation of the multi-omics data.

The genome was then submitted to NCBI as a Whole Genome Shotgun (WGS) project for annotation using the Prokaryotic Genomes Annotation Pipeline (Haft *et al.*, 2018) under BioProject PRJNA473075 and Biosample SAMN09259304, for a total of 3,291 predicted protein coding sequences. A search for 16S RNA gene sequences in the *H. methylovorum* Bras1 2018 draft genome on the basis of NCBI and KEGG orthology has identified one copy of the 16S RNA gene with the accession number DLM45_09350. A BLASTn search for this 1,484 bp sequence against the NCBI nucleotide collection returns the 16S RNA genes of 100% sequence identity (E-value 0.0) to *H. methylovorum* strain NBRC 14180 (NCBI Accession NR_113655.1). A complete copy of the 16S RNA gene sequence and its alignment to the NBRC 16S sequence is shown in Supplementary Figure 5.3-1.

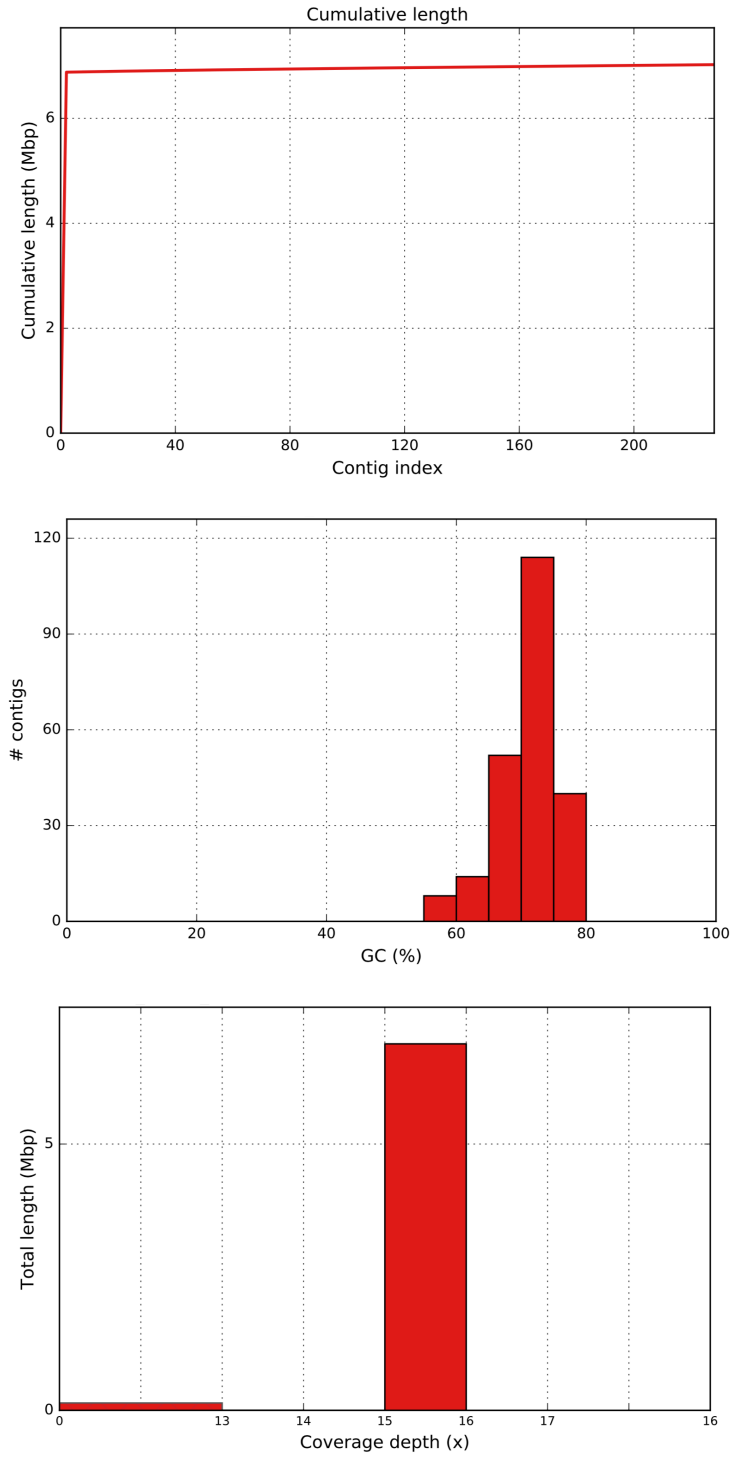


Figure 5.3-1: *Statistical plots of the H. methylovorum Bras1 2018 draft genome pre-removal of short contigs (<1000 bp).* (TOP) Cumulative length the genome represented by the minimum number of contigs. (MIDDLE) Histogram displaying the number of contigs with a given GC content (%) in brackets of 10%. (BOTTOM) Histogram displaying the average contig length (kbp) coverage depth within a given bracket of coverage depth (x). Statistical plots have been generated using the Quast genome quality assessment tool (Gurevich *et al.*, 2013). Note that nucleotide length is doubled due to assessment of forward and reverse sequences.

5.4 Genome analysis of MSC metabolism in model *Hyphomicrobium* species

Functional genomics was previously used to propose metabolic pathways for MSC metabolism in *H. sulfonivorans* S1 (Chapter 3), helping to place proteomic and transcriptomic data into the context of MSC utilisation by the organism (Chapter 4). This section details the metabolic pathways of MSC metabolism in *H. denitrificans* ATCC 51888 and *H. methylovorum* Bras1. For brevity and to avoid redundancy with the proposed pathways of MSC metabolism in *H. sulfonivorans* S1, a greater emphasis has been placed on a comparison between *Hyphomicrobium* species rather than a detailed explanation of each gene assignment, with the exception of genes and enzyme systems of particular interest. This has been divided into four sections: MeOH metabolism (Section 5.4.1), MSC metabolism (Section 5.4.2), formaldehyde metabolism (Section 5.4.3) and inorganic sulfur compound metabolism (Section 5.4.4). Note that BLAST searches (tblastn) were performed using the NCBI BLAST tool (Boratyn *et al.*, 2013), while gene annotations were performed using the NCBI Prokaryotic Genomes Annotation Pipeline (Haft *et al.*, 2018) and functional KEGG BlastKOALA annotation (Kanehisa *et al.*, 2016).

5.4.1 Methanol Metabolism

Genome analysis of *H. denitrificans* suggests that the organism contains an MxaFI-type MeOH dehydrogenase encoded by an *mxaFJIG* gene cluster, plus three XoxF-type MeOH dehydrogenases based on their strong homology to the MxaF-type and XoxF-type methanol dehydrogenases of *Methylobacterium extorquens* (Schmidt *et al.*, 2010; Nakagawa *et al.*, 2012).

Similarly, the *H. methylovorum* genome also contains a putative MxaFI-type MeOH dehydrogenase in a *mxaFJIG* cluster, plus two relatively strong homologues of XoxF-type MeOH dehydrogenases. The results of these BLAST searches can be found in Supplementary Tables S5.4-1 and S5.4-2. Genome analysis of MeOH metabolism in *H. denitrificans* and *H. methylovorum* have been used to construct a metabolic pathway for the organism, which are displayed in Figures 5.4-1 and 5.4-2 respectively.

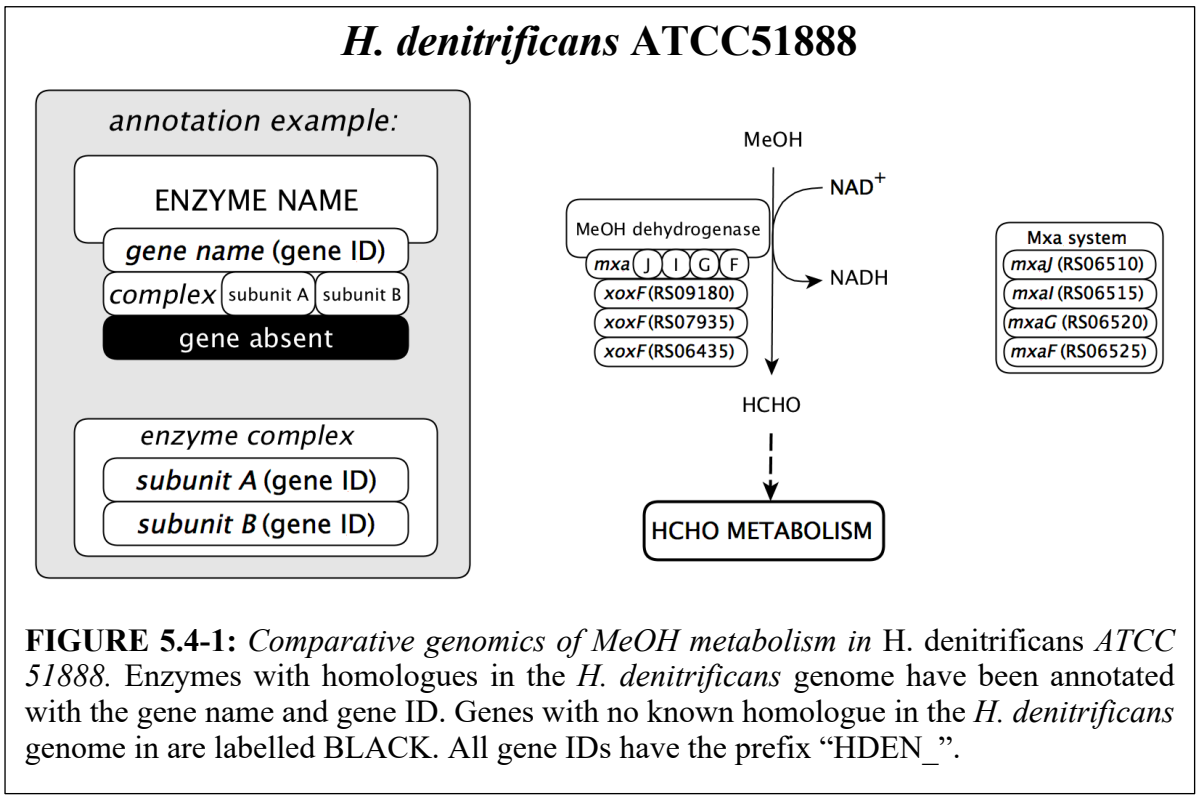


FIGURE 5.4-1: *Comparative genomics of MeOH metabolism in* H. denitrificans *ATCC 51888.* Enzymes with homologues in the *H. denitrificans* genome have been annotated with the gene name and gene ID. Genes with no known homologue in the *H. denitrificans* genome in are labelled BLACK. All gene IDs have the prefix "HDEN_".

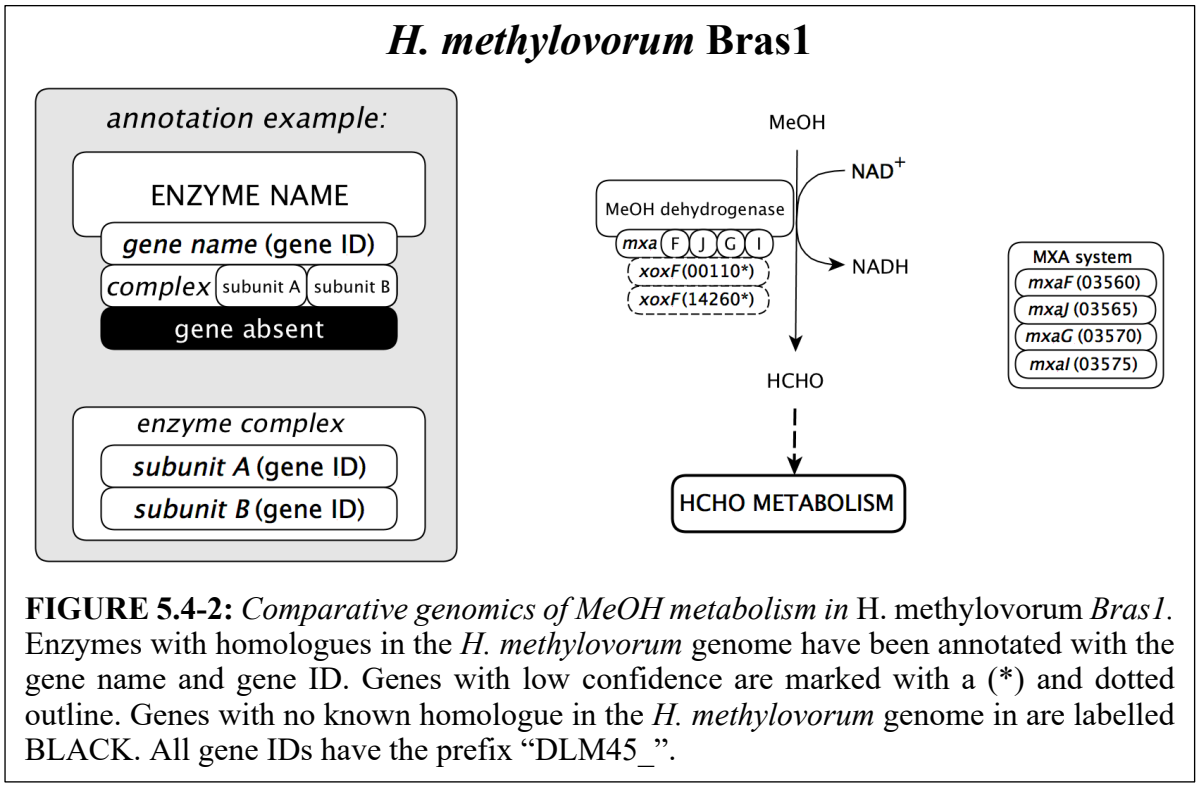


FIGURE 5.4-2: *Comparative genomics of MeOH metabolism in* H. methylovorum *Bras1.* Enzymes with homologues in the *H. methylovorum* genome have been annotated with the gene name and gene ID. Genes with low confidence are marked with a (*) and dotted outline. Genes with no known homologue in the *H. methylovorum* genome in are labelled BLACK. All gene IDs have the prefix "DLM45_".

5.4.2 Methylated Sulfur Compound Metabolism

MSC metabolism in H. denitrificans *ATCC 51888*

*Hyphomicrobiu*m phenotyping in Section 5.2 showed that *H. denitrificans* is capable of utilising DMS as a sole carbon and sole sulfur source. Although *H. denitrificans* lacks a clear homologue of the DmoA-type DMS monooxygenase, it does contain a close homologue of the MtoX-type MT oxidase from *H.* sp. VS (Eyice *et al.*, 2017); HDEN_RS03620.

A search for the $FMNH_2$-dependent monooxygenases of the $DMSO_2$ oxidation pathway has failed to find significant homologues, nor does the organism contain a homologue of the MsmABCD-type MSA monooxygenase system. Indeed, the only $FMNH_2$-dependent monooxygenase with even poor homology to any of these enzymes is HDEN_RS08615, annotated by NCBI as a putative flavin-dependent oxidoreductase with no function predicted by KEGG orthology. This lack of $FMNH_2$-dependent monooxygenases starkly contrasts with the numerous enzymes previously described for *H. sulfonivorans* in Chapters 3 and 4, which appear to play a major role $DMSO_2$ metabolism by the organism.

A tentative map of MSC metabolism in *H. denitrificans* for functional genomics (See Section 5.5) is displayed in Figure 5.4-3, while the BLAST results of MSC metabolism in the organism can be found in Supplementary Table S5.4-3.

MSC metabolism in H. methylovorum *Bras1*

H. methylovorum appears to have close homologues for both the DmoA-type DMS monooxygenase (DLM45_02930) and MtoX-type MT oxidase (DLM45_12340), making it the only one of the three model organisms to contain both enzymes of the putative DMS oxidation pathway.

For the $DMSO_2$ oxidation pathway, proposed as an alternate pathway of $DMSO_2$ metabolism in *H. sulfonivorans*, two potential enzymes have also been found. An SfnFG-type $DMSO_2$ monooxygenase appears to be present, encoded by the putative FMN-reductase *sfnF* (DLM45_02940) and $FMNH_2$-dependent monooxygenase *sfnG* (DLM45_02945), supported by NCBI annotation and the respective KEGG orthologies K00299 and K17228. Although no MsuDE-type MSA monooxygenase was found in the organism, strong homologues for an MsmABCD-type MSA monooxygenase system have been found in *H. methylovorum* as a series of adjacent genes within the genome (Marco *et al.* 1999).

A tentative map of MSC metabolism in *H. methylovorum* for functional genomics (See Section 5.5) is displayed in Figure 5.4-4, while the BLAST results of MSC metabolism in the organism can be found in Supplementary Table S5.4-4.

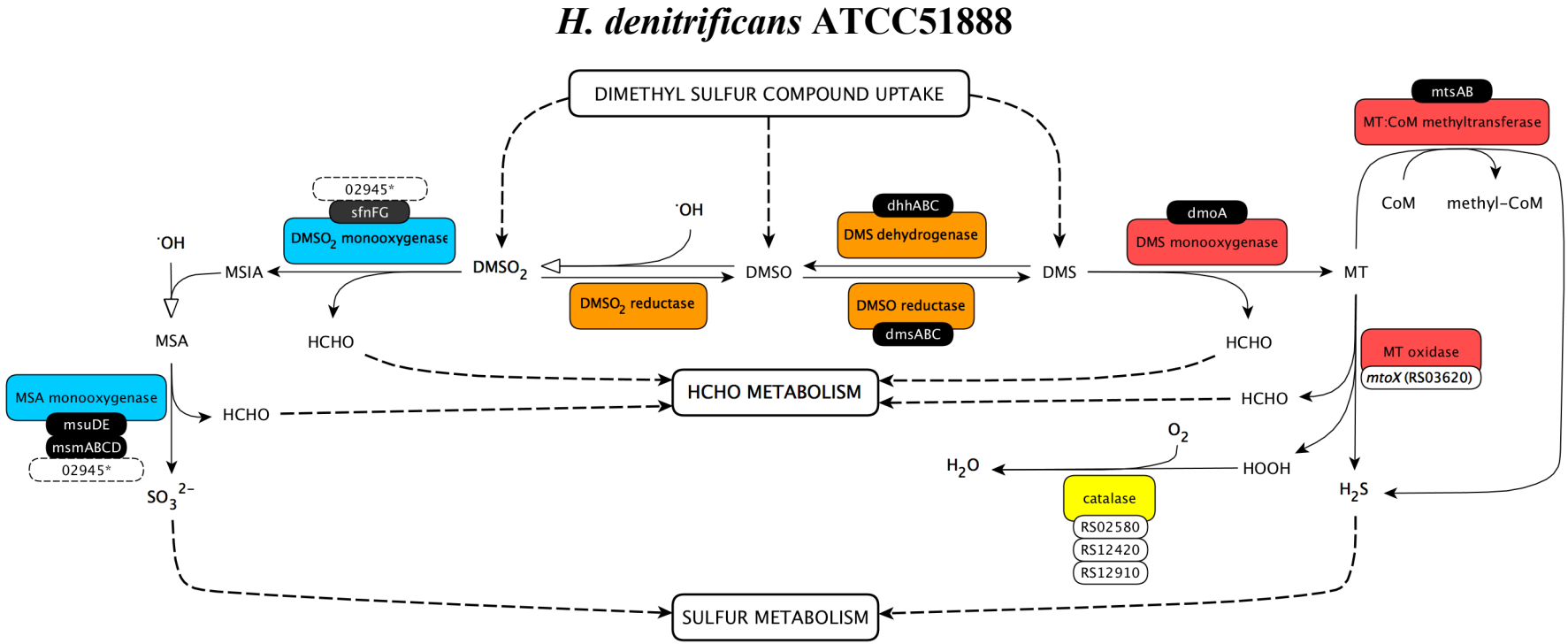


Figure 5.4-3: *Comparative genomics analysis of MSC metabolism in* H. denitrificans *ATCC 51888.* Enzymes with homologues in the *H. denitrificans* genome have been annotated with a WHITE label, containing the gene name and ID. Genes with low confidence are marked with a (*) and dotted outline. Genes with no known homologue in the *H. denitrificans* genome are labelled in BLACK. Mechanisms of MSC metabolism are labelled as follows: Interconversion of DMSCs (ORANGE), DMS oxidation pathway (RED), $DMSO_2$ oxidation pathway (BLUE), hydrogen peroxide detoxification (YELLOW). Non-enzymatic reactions are marked by a WHITE arrowhead.

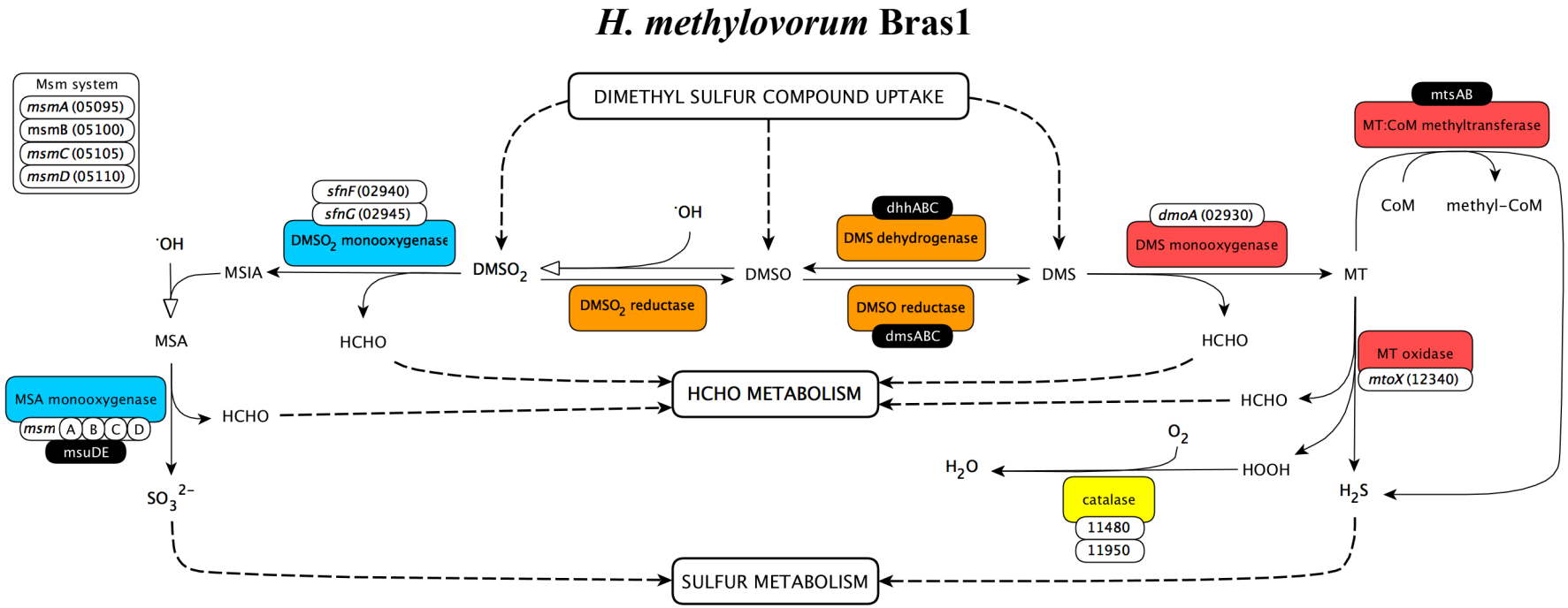


Figure 5.4-4: *Comparative genomics analysis of MSC metabolism in* H. methylovorum *Bras1.* Enzymes with homologues in the *H. methylovorum* genome have been annotated with a WHITE label, containing the gene name and ID. Genes with low confidence are marked with a (*) and dotted outline. Genes with no known homologue in the *H. methylovorum* genome are labelled in BLACK. Mechanisms of MSC metabolism are labelled as follows: Interconversion of DMSCs (ORANGE), DMS oxidation pathway (RED), $DMSO_2$ oxidation pathway (BLUE), hydrogen peroxide detoxification (YELLOW). Non-enzymatic reactions are marked by a WHITE arrowhead.

5.4.3 Formaldehyde Metabolism

As serine cycle methylotrophs, methylotrophic formaldehyde assimilation in *Hyphomicrobium* species is likely to occur via the production and assimilation of 5,10-methylene H_4F, potentially mediated by a combination of the H_4MPT-dependent formaldehyde activation pathway and H_4F pathway, as previously described for *M. extorquens* (Crowther *et al.*, 2008). However, some evidence also exists for the presence of formaldehyde and formate dehydrogenases, with the latter involved in the dissimilation of formate to CO_2.

Formaldehyde metabolism in H. denitrificans ATCC 51888

A search of the *H. denitrificans* genome for homologues of known enzymes of formaldehyde metabolism has yielded candidates for a complete H_4MPT-dependent formaldehyde activation pathway, included as many as three copies of an Fae-type formaldehyde activating enzyme and a single copy of the MtdB-like methylene H_4MPT dehydrogenase, Mch-like methenyl H_4MPT cyclohydrolase and the bifunctional FhcABCD-type 5-formyl H_4MPT N-formyltransferase/ formyl MFR dehydrogenase.

In terms of formate metabolism, *H. denitrificans* has a single Fdh-type formate dehydrogenase homologue but two copies of the H_4F pathway proposed to mediate the production of 5,10-methylene H_4F for carbon assimilation, Fhs-type formate H_4F ligases and a FolD-type bifunctional methenyl H_4F cyclohydrolase/ methylene H_4F dehydrogenase. Note that while two of these genes are localised to a *folD/fch* cluster (HDEN_RS00520/HDEN_RS00515) the other *folD* (HDEN_RS15885) and *fch* (HDEN_RS15000) are located elsewhere in the genome.

A tentative map of formaldehyde metabolism in *H. denitrificans* is displayed in Figure 5.4-5, while the BLAST results of formaldehyde metabolism in the organism can be found in Supplementary Table S5.4-5.

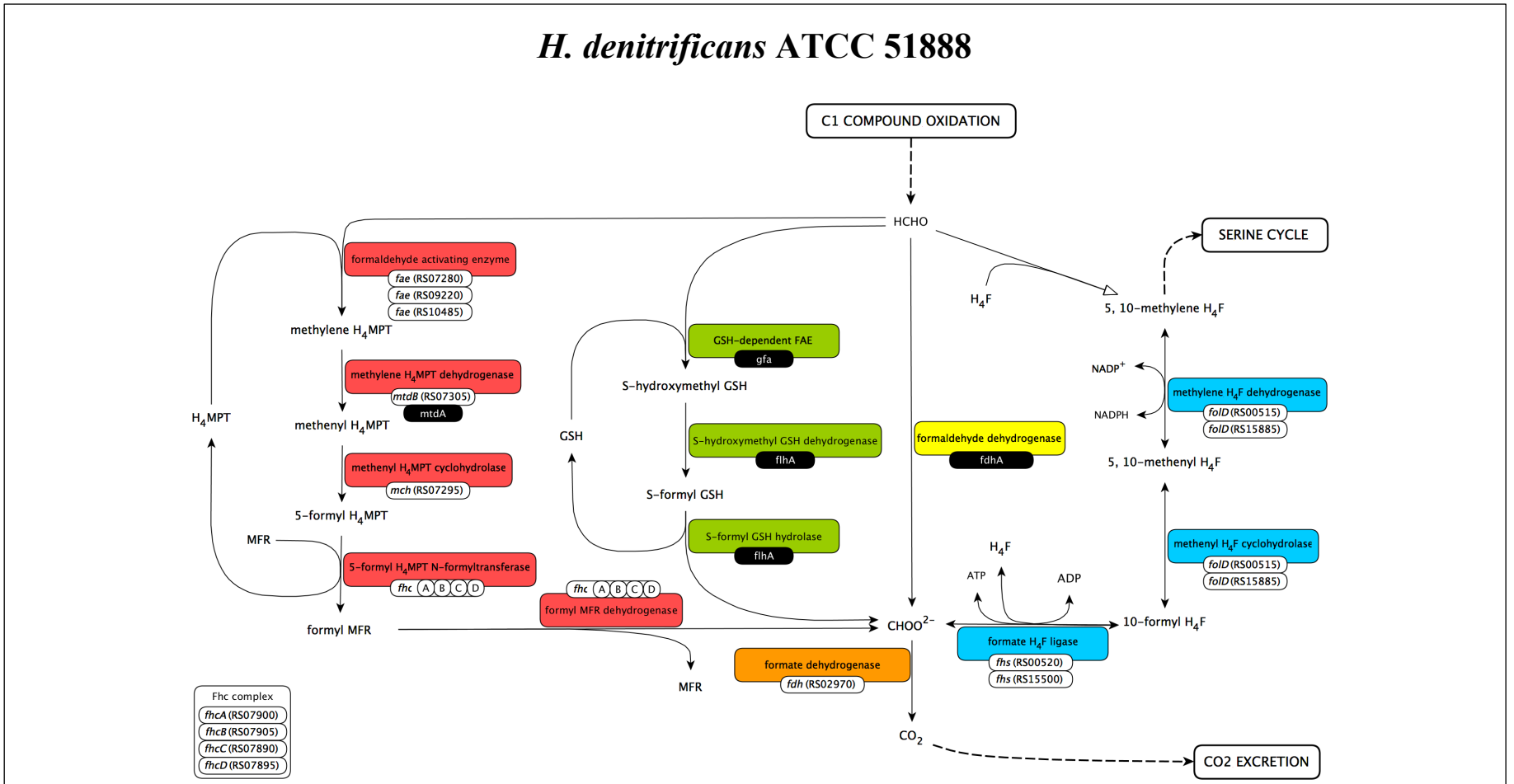


Figure 5.4-5: *Comparative genomics analysis of formaldehyde metabolism in* H. denitrificans *ATCC 51888.* Enzymes with homologues in the *H. denitrificans* genome have been annotated with a WHITE label, containing the gene name and ID. Genes with no known homologue in the *H. denitrificans* genome are labelled in BLACK. Mechanisms of formaldehyde (HCHO) metabolism are labelled as follows: H_4MPT-dependent pathway (RED), GSH-dependent pathway (GREEN), formaldehyde oxidation (YELLOW), formate oxidation (ORANGE) and H_4F pathway (BLUE). Non-enzymatic reactions are marked by a WHITE arrowhead.

Formaldehyde metabolism in H. methylovorum *Bras1*

Formaldehyde metabolism in *H. methylovorum* appears to be much the same as that of *H. denitrificans*, with an H_4MPT-dependent formaldehyde activation pathway, but no convincing candidates for either a GSH-dependent pathway or formaldehyde dehydrogenase, with the closest candidate for the enzymes of both systems being a poor homology alcohol dehydrogenase (DLM45_11985). The only difference being that *H. methylovorum* has two copies of the Fae-type formaldehyde activating enzyme rather than three. Although *H. methylovorum* contains no clear Fdh-type formate dehydrogenase, it does contain two copies of the H_4F pathway enzymes H_4F ligases (DLM45_07050 and DLM45_13340) and FolD-type bifunctional methenyl H_4F cyclohydrolase/ methylene H_4F dehydrogenase (DLM45_07540 and DLM45_13300), localised within two broad gene clusters each containing an *fch* and *folD*.

A tentative map of formaldehyde metabolism in *H. methylovorum* is displayed in Figure 5.4-6, while the formaldehyde metabolism BLAST results for the organism can be found in Supplementary Table S5.4-6.

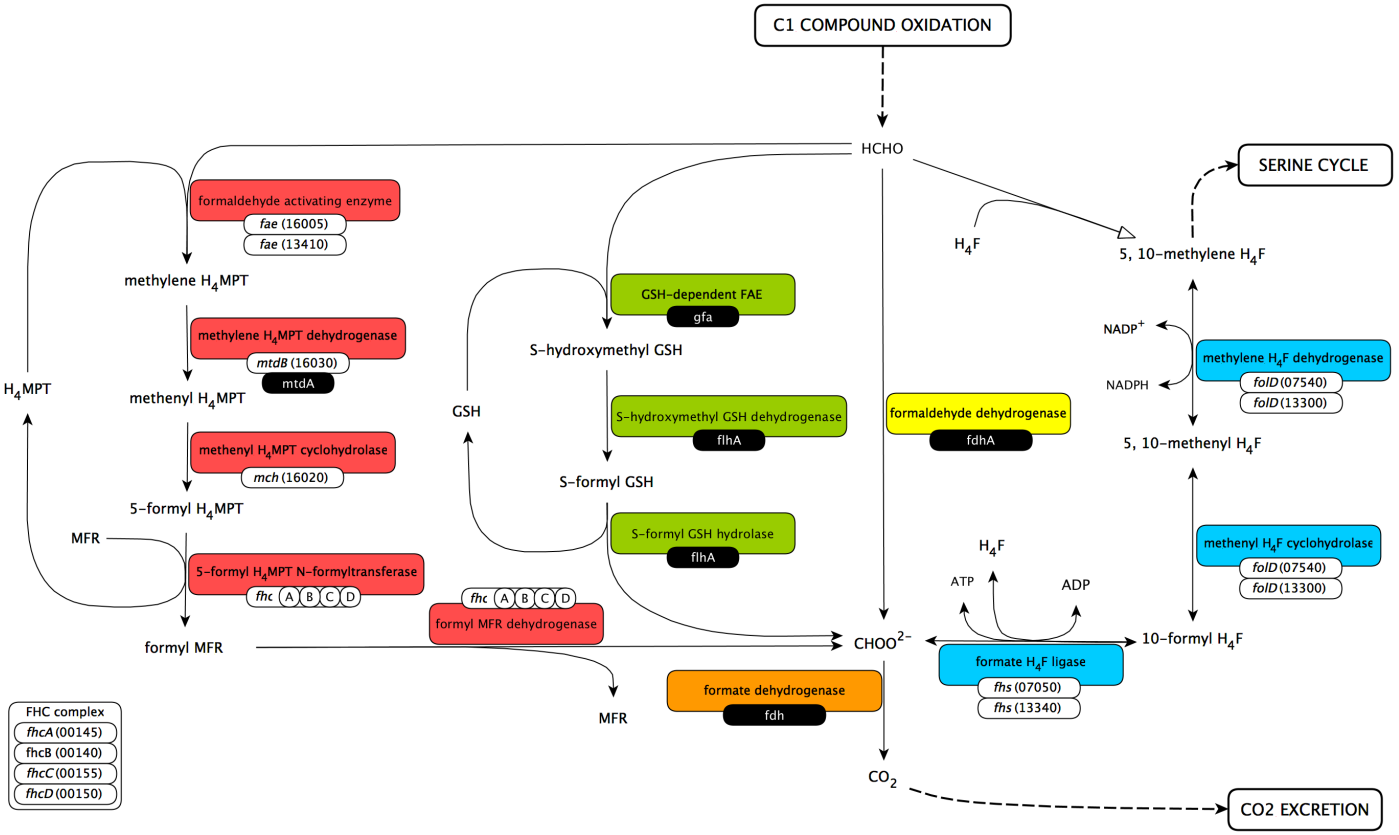


Figure 5.4-6: *Comparative genomics analysis of formaldehyde metabolism in* H. methylovorum *Bras1.* Enzymes with homologues in the *H. methylovorum* genome have been annotated with a WHITE label, containing the gene name and ID. Genes with low confidence are marked with a (*) and dotted outline. Genes with no known homologue in the *H. methylovorum* genome are in labelled BLACK. Mechanisms of formaldehyde (HCHO) metabolism are labelled as follows: H_4MPT-dependent pathway (RED), GSH-dependent pathway (GREEN), formaldehyde oxidation (YELLOW), formate oxidation (ORANGE) and H_4F pathway (BLUE). Non-enzymatic reactions are marked by a WHITE arrowhead.

5.4.4 Inorganic Sulfur Metabolism

A genomic analysis of inorganic sulfur compound metabolism in *H. denitrificans* and *H. methylovorum* is outlined below, split between general sulfur metabolism and an analysis of the Sox system due to the wide array of Sox genes present in both species.

Inorganic sulfur compound metabolism in H. denitrificans ATCC 51888

Based on the genomics of MSC metabolism in *H. denitrificans* discussed above, it is probable that DMS metabolism yields hydrogen sulfide but unlikely to produce sulfite due to its lack of a $DMSO_2$ oxidation pathway.

Koch and Dahl (2018) have proposed that sulfite oxidation in *H. denitrificans* ATCC 51888 may occur via the non-enzymatic production of thiosulfate from sulfite and hydrogen sulfide, followed by the Hdr-mediated oxidation of thiosulfate to generate sulfite. Each of the major enzymes of this process have been successfully identified in the genome annotation of *H. denitrificans* used for this project, including the HdrAB1B2C1C2-type heterodisulfide reductase encoded by *hdrA*, *hdrB1*, *hdrB2*, *hdrC1* and *hdrC2*, as well as the associated LbpA-type lipoate-binding protein and TusA-type sulfur carrier protein *tusA*.

Although assimilatory sulfate metabolism in *H. sulfonivorans* is likely to occur via the oxidation of sulfate to sulfite by a sulfate adenylylation pathway, only a partial pathway has been identified in the *H. denitrificans* genome. Homologues have been found for a cysNCD-type bifunctional sulfate adenylyltransferase/ adenylyl-sulfate kinase, consisting of the CysNC (HDEN_RS05165) and CysD (HDEN_RS05170) subunits, and a CysH-type phosphoadenylyl-sulfate reductase (HDEN_RS05175).

A tentative map of inorganic sulfur compound metabolism in *H. denitrificans* for functional genomics (See Section 5.5) is displayed in Figure 5.4-6.

The Sox system of H. denitrificans ATCC 51888

H. denitrificans has homologues of the entire SoxAXBCDYZ complex split between several gene clusters and may exhibit some redundancy between Sox enzymes. These Sox genes have been identified by performing BLAST searches for the experimentally characterised Sox genes of *Paracoccus pantotrophus* and the putative SoxAXBCYZ genes of *H. denitrificans* ATCC 51888 recently proposed by Koch and Dahl (2018). Note that while the reference genome for these Sox sequences is the same used by Koch and Dahl (2018) this

project has instead used the NCBI annotation of the *H. denitrificans* genome for consistency, so there will be some differences in the score and E-value of these BLAST results.

The largest cluster of Sox genes in *H. denitrificans* is a SoxAXBYZ, identified by Koch and Dahl (2018) and encoded by a *soxA*, *soxX*, *soxB*, *soxY* and *soxZ*. A smaller SoxYZCD cluster is also present elsewhere in the genome, consisting of SoxYZ-like carrier encoding *soxYZ* and a putative SoxCD encoded by a *soxC* and a *soxD* gene. Moreover *H. denitrificans* also contains two more SoxYZ-like carriers (*soxYZ*) and an additional SoxYZ (*soxY*, *soxZ*). Finally, the organism also contains three more SoxD homologues (HDEN_RS07325, HDEN_RS10760, HDEN_RS13510), which have been tentatively labelled as potential *soxD* genes.

Inorganic sulfur compound metabolism in H. methylovorum Bras1

Like *H. sulfonivorans*, *H. methylovorum* has both a putative DMS oxidation pathway and a putative $DMSO_2$ oxidation pathway, suggesting that MSC utilisation by the organism may produce hydrogen sulfide or sulfite. Beginning with sulfite metabolism, the organism contains a good homologue for a CysJ-type sulfite reductase but no obvious sulfite dehydrogenase, consistent with the other two *Hyphomicrobium* species. For inorganic sulfur assimilation, the organism has two genes encoding MetZ-like homocysteine synthases and one CysK-like cysteine synthase A, plus the CysE and Acs enzymes necessary to generate precursors and cofactors for cysteine synthesis.

H. methylovorum lacks any significant homologues for the various enzymes of Hdr-mediated thiosulfate oxidation characterised in *H. denitrificans*, as neither HdrABC, TusA nor LbpA. Finally, two enzymes of the sulfate adenylylation pathway have been identified in *H. methylovorum*, providing a mechanism for the oxidation of sulfate to sulfite for use as an inorganic sulfur source. These are a bifunctional CysNCD-type sulfate adenylyltransferase/ adenylyl-sulfate kinase, consisting of a CysNC and CysD subunit, and a cysH-type phosphoadenylyl-sulfate reductase. This complement differs slightly from the pathway of *H. sulfonivorans* in that it lacks an additional CysC.

A tentative map of inorganic sulfur compound metabolism in *H. methylovorum* for functional genomics (See Section 5.5) is displayed in Figure 5.4-6.

The SOX system of H. methylovorum Bras1

Surprisingly, *H. methylovorum* shows even more homologues of Sox system enzymes than *H. denitrificans*, with two complete SoxAXBCDYZ clusters plus several other Sox genes. Based on localisation and homology to the *H. denitrificans* Sox gene sequences identified by Koch and Dahl (2018), the majority of Sox genes can be divided into two clusters. SoxAXBCDYZ DLM45_12160:12195 cluster displays the highest homology to the *H. denitrificans* reference sequences and is split into two outward-facing series of the five Sox genes *soxYZBCD* and the two Sox genes *soxAX*.

The second SoxAXBCDYZ DLM45_14665-14695 cluster consists of seven adjacent Sox genes *soxXYZABCD*, all in the same orientation but showing relatively poor homology to the Sox enzymes of the first cluster. For example, while the SoxA (DLM45_12165) and SoxX (DLM45_12160) of the first cluster show 72% and 54% sequence identity to their respective homologues in *H. denitrificans*, the second cluster's putative SoxA (DLM45_14680) and SoxX (DLM45_14665) only exhibit a 30% and nil (E-value >0.05) sequence identity to the same enzymes.

In addition to these two SoxAXBCDYZ encoding gene clusters, *H. methylovorum* also contains various other Sox genes encoding a SoxCD (*soxC, soxD*), SoxYZ (*soxY*, *soxZ*), two SoxYZ-like carrier proteins (*soxYZ*) and a low homology SoxD-like cytochrome *c*. As *H. methylovorum* has a multitude of Sox enzymes it is unknown how many of these putative Sox genes are responsible for mediating thiosulfate oxidation, so all of them have been added to the organism's metabolic pathway to examine their response to MSCs using comparative proteomics.

The presence of two distinct Sox gene clusters raised the concerning possibility that the *H. methylovorum* sequence data had been contaminated by genomic DNA from another organism, especially as the Bras1 2018 genome is a hybrid assembly from two different Illumina sequencing experiments; one set of sequence data originated from this project and the other coming from a previous experiment performed by the Schäfer lab group (unpublished). However, a BLAST search for these enzymes on individual draft assemblies compiled from each set of sequence data indicates that both Sox clusters are present in both sets of Illumina sequence data. This makes it far more likely that this second, low homology DLM45_14665-14695 Sox gene cluster is the product of horizontal gene transfer rather than contamination.

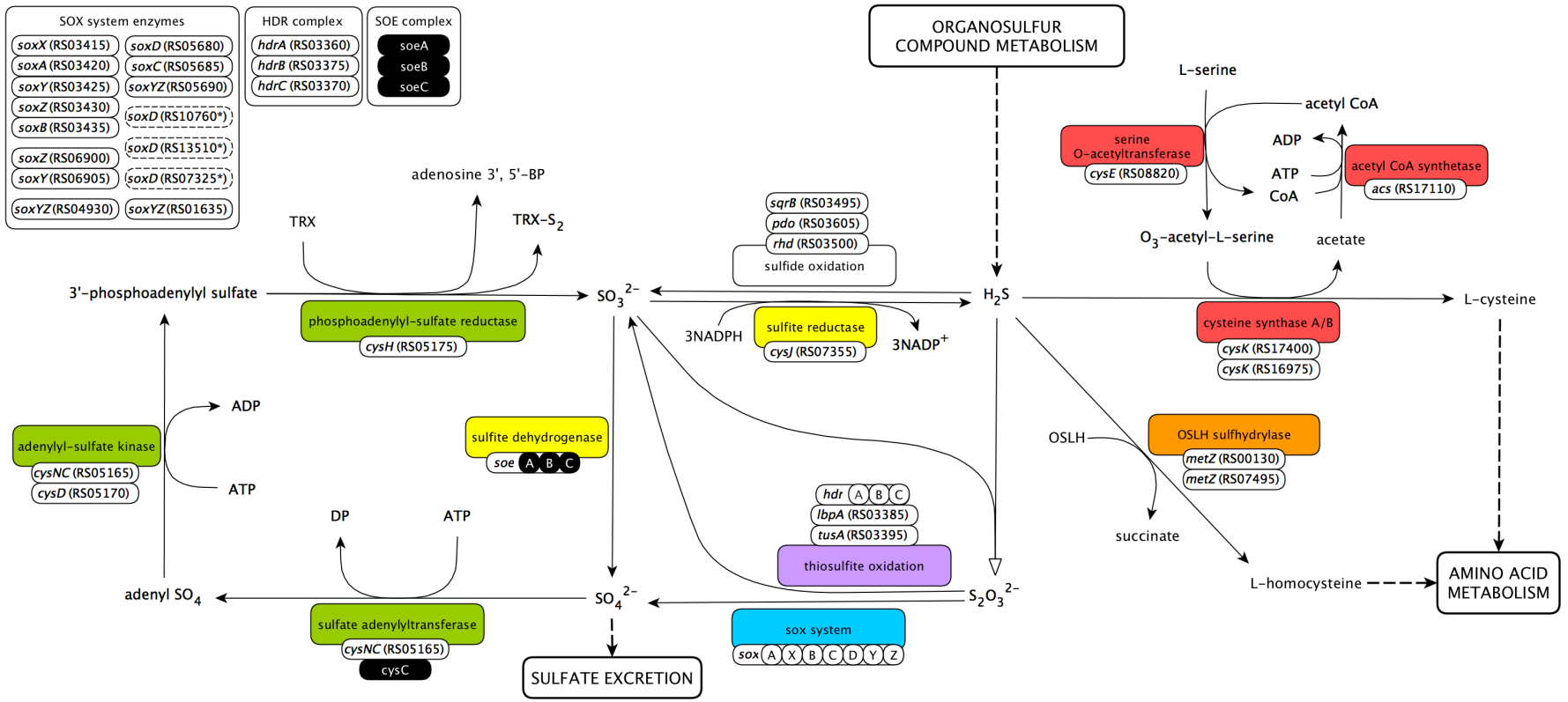


Figure 5.4-7: *Comparative genomics analysis of inorganic sulfur compound MSC metabolism in* H. denitrificans *ATCC 51888.* Enzymes with homologues in the *H. denitrificans* genome have been annotated with a WHITE label, containing the gene name and ID. Genes with low confidence are marked with a (*) and dotted outline. Genes with no known homologue in the *H. denitrificans* genome are labelled BLACK. Mechanisms of inorganic sulfur compound metabolism are labelled as follows: sulfite metabolism (YELLOW), cysteine synthesis (RED), homocysteine synthesis (ORANGE), Sox-mediated thiosulfate oxidation (BLUE), Hdr-mediated thiosulfate oxidation (PURPLE), sulfate adenylylation pathway (GREEN). Non-enzymatic reactions are marked by a WHITE arrowhead.

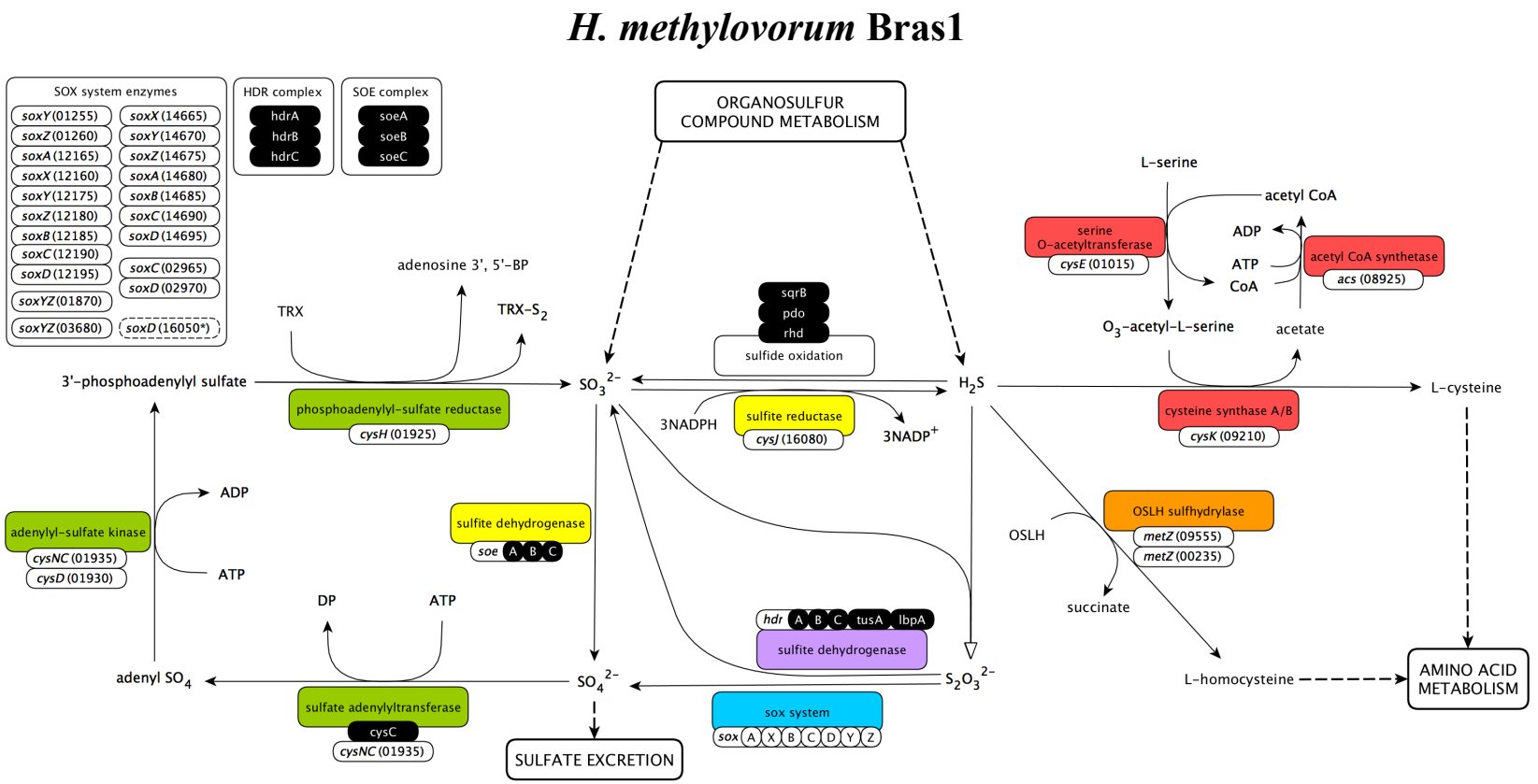


Figure 5.4-8: *Comparative genomics analysis of inorganic sulfur compound MSC metabolism in* H. methylovorum *Bras1.* Enzymes with homologues in the *H. methylovorum* genome have been annotated with a WHITE label, containing the gene name and ID. Genes with low confidence are marked with a (*) and dotted outline. Genes with no known homologue in the *H. methylovorum* genome are labelled BLACK. Mechanisms of inorganic sulfur compound metabolism are labelled as follows: sulfite metabolism (YELLOW), cysteine synthesis (RED), homocysteine synthesis (ORANGE), Sox-mediated thiosulfate oxidation (BLUE), Hdr-mediated thiosulfate oxidation (PURPLE), sulfate adenylylation pathway (GREEN). Non-enzymatic reactions are marked by a WHITE arrowhead.

5.4.5 Conclusions

Although the metabolic pathways of formaldehyde and inorganic sulfur compound metabolism that have been assembled for *H. denitrificans* and *H. methylovorum* appear to be consistent with that of *H. sulfonivorans*, these species contain a disparate array of MSC oxidases and monooxygenases.

The *H. sulfonivorans* genome contains multiple $FMNH_2$-dependent monooxygenases, including a DmoAB monooxygenase, SfnFG-type $DMSO_2$ monooxygenase, MsuDE-type MSA monooxygenase and another SfnFG-like monooxygenase of unknown function. In contrast, *H. methylovorum* contains a putative DmoAB monooxygenase, SfnFG-type monooxygenase, MtoX-type MT oxidase and an MsmABCD-type MSA monooxygenase, while *H. denitrificans* only contains an MtoX oxidase. This suggests that *H. sulfonivorans* and *H. methylovorum* contain both a $DMSO_2$ oxidation pathway and a DMS oxidation pathway, while *H. denitrificans* only contains a DMS oxidation pathway,

As these pathways are predicted to generate different sulfur products, it may be unsurprising that key differences also exist between the several sulfur oxidation systems of each organism. Most notably, that the Hdr-type thiosulfate oxidation system recently identified in *H. denitrificans* is missing from both *H. sulfonivorans* and *H. methylovorum*, and that *H. methylovorum* contains two copies of the Sox system while *H. sulfonivorans* only has a partial Sox system. It is unknown what effect such variation will have on these organism's capacity to utilise inorganic sulfur compounds, or indeed the methylotrophic utilisation of MSCs which is predicted to produce a large excess of inorganic sulfur.

It will therefore be interesting to see how these distinct sets of metabolic enzymes interact with each other using comparative proteomics in response to different MSCs.

5.5 Comparative Proteomics of MSC utilisation in *Hyphomicrobium species*

The phenotyping of *Hyphomicrobium* species performed in Section 5.2 demonstrated that different *Hyphomicrobium* species are capable of utilising a different array of MSCs as sole carbon or sole sulfur source. One of the greatest differences is in their ability, or inability, to utilise $DMSO_2$ or DMS as a sole carbon source, with *H. denitrificans* utilising DMS, *H. sulfonivorans* utilising $DMSO_2$ and *H. methylovorum* utilising both DMS and $DMSO_2$. Genome analysis of MSC metabolism in these species shows that this difference in phenotype correlates to the presences or absence of two enzymes; a DMS monooxygenase is present in each species that can utilise $DMSO_2$ and a MT oxidise is present in the species that can utilise DMS. To better understand this relationship between genotype and phenotype, comparative proteomics has been used to examine MSC metabolism in *H. denitrificans* ATCC 51888 and *H. methylovorum* Bras1.

5.5.1 Aims and experimental design

The primary aim is to use comparative proteomics to compare and contrast the utilisation of DMS and $DMSO_2$ as a carbon source in *Hyphomicrobium* species. By examining MSC utilisation in one species that can utilise $DMSO_2$, one that can utilise DMS and one that can utilise both $DMSO_2$ and DMS, we hope to better understand the molecular mechanisms that each *Hyphomicrobium* species uses to degrade different MSCs and discover an explanation for why each species is able or unable to grow on these MSCs.

Comparative proteomics has already been performed for the growth of *H. sulfonivorans* on $DMSO_2$, so new proteomics data only needed to be collected for *H. denitrificans* on DMS and *H. methylovorum* on DMS and $DMSO_2$. This was performed in a separate proteomics experiment for each species, largely following the work flow of the previous proteomics experiment for *H. sulfonivorans*: Bacterial biomass was generated by fed-batch cultivation, protein was extracted and digested to generate peptides, peptides were analysed by mass spectrometry and the data used to create a proteome for each condition (see Section 4.3).

Comparative Proteomics of Hyphomicrobium denitrificans ATCC 51888

Comparative proteomics was performed to examine the metabolism of DMS by *H. denitrificans* ATCC 51888 when cultivated on a sole carbon source of DMS (DMS condition) against a control condition cultivated on a carbon source of MeOH (Control condition). All cultures were cultivated on defined, carbon-limited media supplemented with sulfate as a replete inorganic sulfur source, with additional carbon source added to the culture over time. An overview of this experiment is shown in Figure 5.5-1.

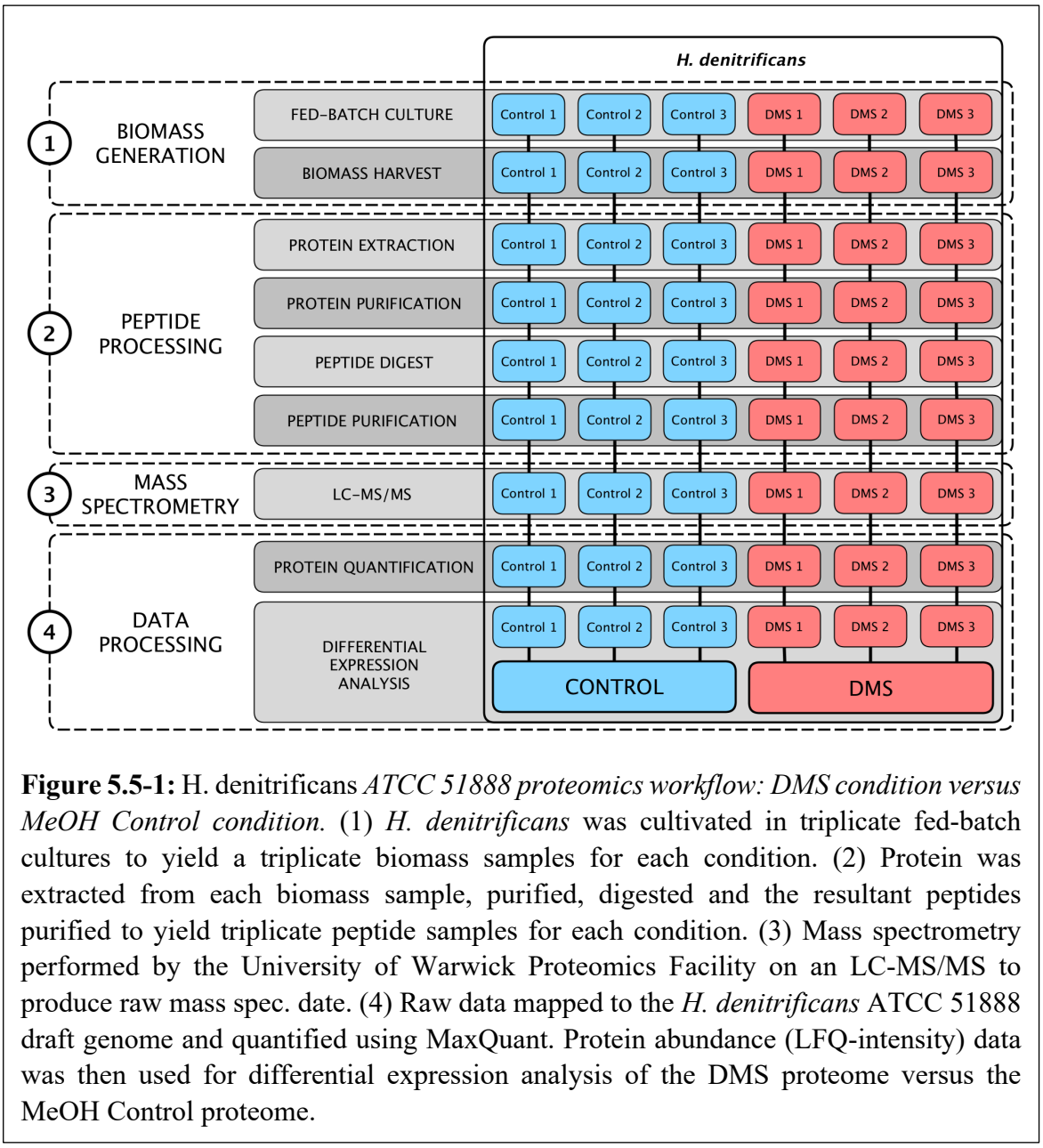


Figure 5.5-1: H. denitrificans *ATCC 51888 proteomics workflow: DMS condition versus MeOH Control condition.* (1) *H. denitrificans* was cultivated in triplicate fed-batch cultures to yield a triplicate biomass samples for each condition. (2) Protein was extracted from each biomass sample, purified, digested and the resultant peptides purified to yield triplicate peptide samples for each condition. (3) Mass spectrometry performed by the University of Warwick Proteomics Facility on an LC-MS/MS to produce raw mass spec. date. (4) Raw data mapped to the *H. denitrificans* ATCC 51888 draft genome and quantified using MaxQuant. Protein abundance (LFQ-intensity) data was then used for differential expression analysis of the DMS proteome versus the MeOH Control proteome.

Each condition was cultivated in triplicate for 244 hours in fed-batch for a total of six cultures, supplemented with additional carbon source at 72, 120, 168, 216, 228, 240 and 243 hrs (see Figure 5.5-2). At 244 hrs bacterial biomass was harvested from each culture, protein extracted, the extracts purified by SDS-page gel electrophoresis and then digested to produce triplicate peptide samples for each condition. Peptides were then analysed by Liquid Chromatography–Mass Spectrometry/ Mass Spectrometry (LC-MS/MS) and mapped to the *H. denitrificans* ATCC 51888 reference genome for quantification, generating an abundance for each protein (LFQ-intensity) for differential protein analysis.

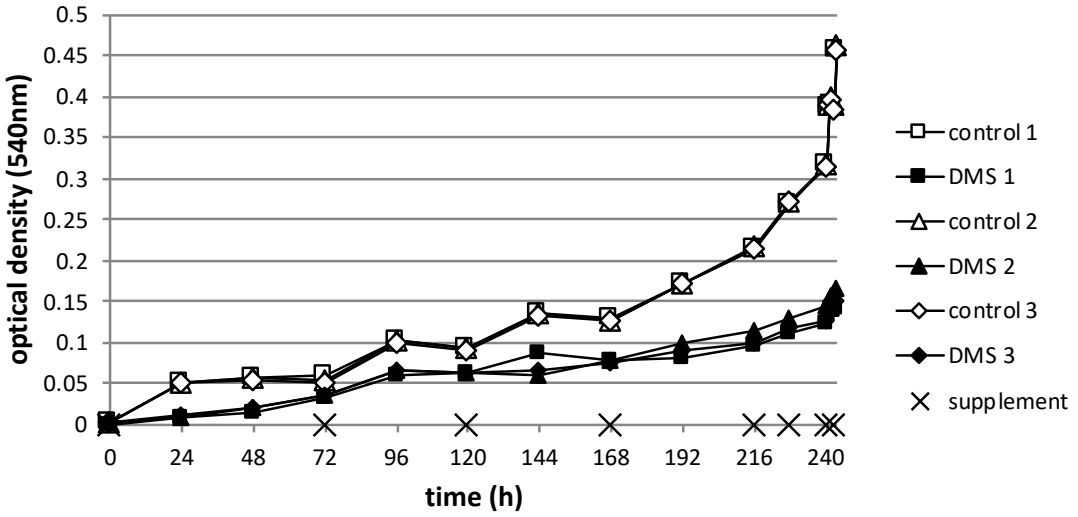


Figure 5.5-2: *Cultivation of* H. denitrificans *for proteomics. H. denitrificans* ATCC 51888 was cultivated in triplicate fed-batch cultures on a MeOH Control condition and a DMS condition. Cultures were supplemented with additional carbon source at regular intervals, marked by an (X).

Comparative Proteomics of Hyphomicrobium methylovorum ATCC 51888

The metabolism of MSCs by *H. methylovorum* Bras1 was examined by comparative proteomics when cultivated on a sole carbon source DMS (DMS condition), sole carbon source of $DMSO_2$ ($DMSO_2$ condition) and a control condition using a sole carbon source of MeOH (Control condition). All cultures were cultivated on defined, carbon-limited media supplemented with sulfate as a replete inorganic sulfur source, with additional carbon source added to the culture over time. An overview of this experiment is shown in Figure 5.5-3.

Each condition was cultivated in triplicate for 360 hours in fed-batch for a total of nine cultures, supplemented with additional carbon source at 72, 120, 168, 216, 264, 288, 312, 366,

348, 357, 358 and 359 hrs (see Figure 5.5-4). At 360 hrs bacterial biomass was harvested from each culture, the protein extracted, the extracts purified by SDS-page gel electrophoresis and then digested to produce triplicate peptide samples for each condition. Peptides were then analysed by Liquid Chromatography–Mass Spectrometry/ Mass Spectrometry (LC-MS/MS) and mapped to the *H. methylovorum* Bras1 reference genome for quantification, generating an abundance for each protein (LFQ-intensity) for differential protein analysis.

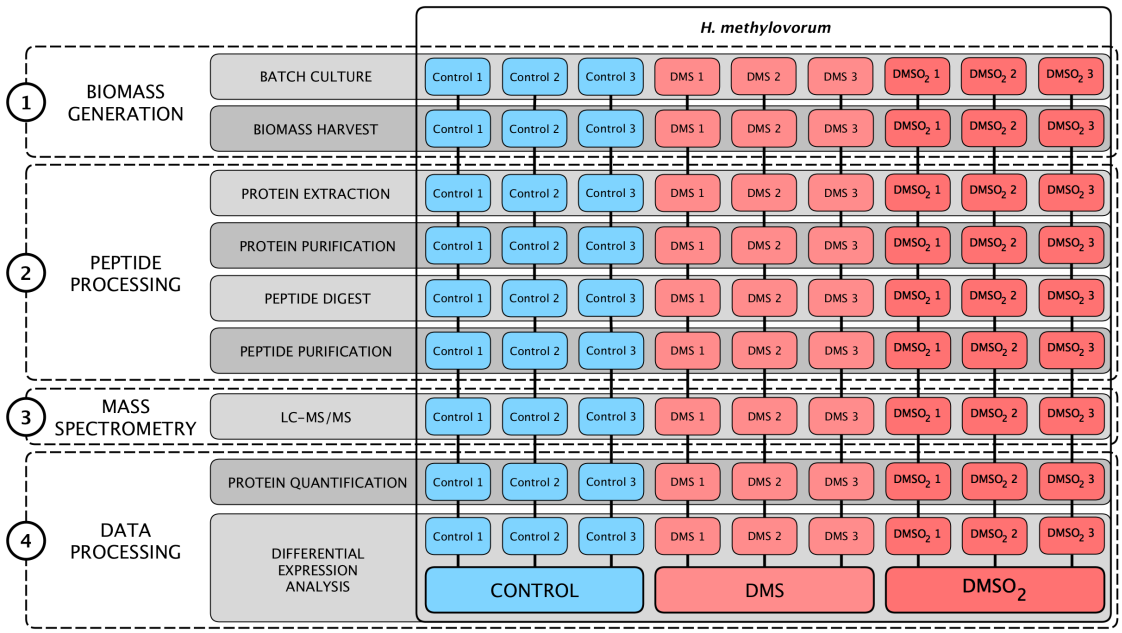


Figure 5.5-3: H. methylovorum *Bras1 proteomics workflow: DMS and $DMSO_2$ conditions versus a MeOH Control condition.* (1) *H. methylovorum* was cultivated in triplicate fed-batch cultures to yield a triplicate biomass samples for each condition. (2) Protein was extracted from each biomass sample, purified, digested and the resultant peptides purified to yield triplicate peptide samples for each condition. (3) Mass spectrometry performed by the University of Warwick Proteomics Facility on an LC-MS/MS to produce raw mass spec. date. (4) Raw data mapped to the *H. methylovorum* Bras1 draft genome and quantified using MaxQuant. Protein abundance (LFQ-intensity) data was then used for differential expression analysis of the DMS proteome versus the MeOH Control proteome.

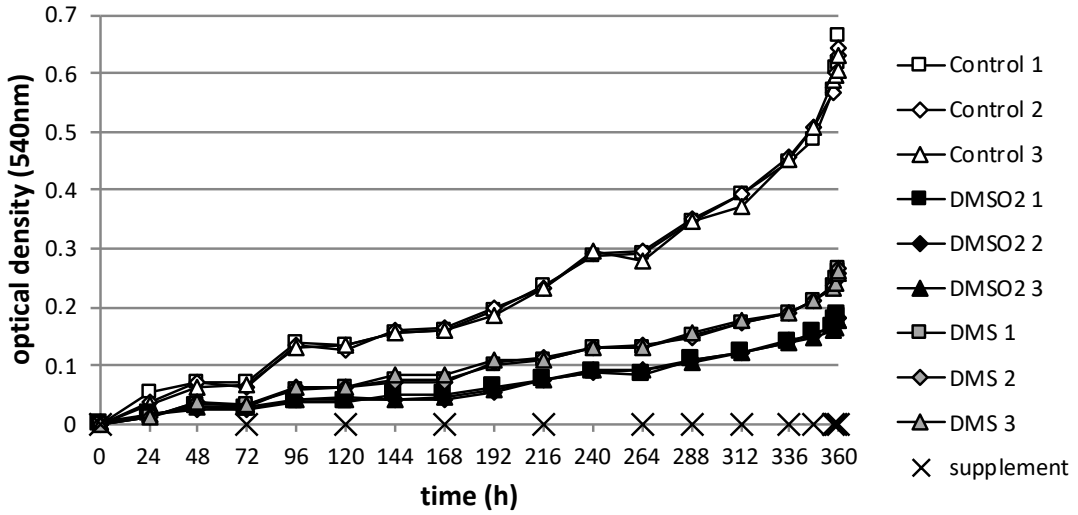


Figure 5.5-4: *Cultivation of* H. methylovorum *for proteomics. H. methylovorum* Bras3 was cultivated in triplicate fed-batch cultures on a MeOH Control condition, a $DMSO_2$ condition and a DMS condition. Cultures were supplemented with additional carbon source at regular intervals, marked by an (X).

5.5.2 Overview of results

The proteomics data sets cover two species, *H. denitrificans* and *H. methylovorum*, cultivated on a sole carbon source of MeOH and DMS, and MeOH, DMS and $DMSO_2$ respectively. Protein abundance data, derived from LFQ-intensity, is expressed as relative abundance of either DMS or $DMSO_2$ compared to the MeOH control condition, i.e. if a protein is upregulated on DMS but downregulated on $DMSO_2$ then it is more abundant on the DMS condition than on the MeOH condition, but less abundant on $DMSO_2$ than the MeOH condition. Note that for a protein to be counted in the data set, it had to be positively identified in three replicates of the same condition, by the presence of at least one unique peptide in the raw mass spectrometry data.

Volcano plots of the *H. denitrificans* and *H. methylovorum* proteomics data are shown in Figure 5.5-5, with the Log2 fold-change of each protein plotted against its -Log p-value, for each pair of test and control conditions. In terms of differential expression and significance, each of the plots show a substantial amount of significant up and downregulation on the MSC conditions versus their MeOH control conditions (q-value <0.05, false discovery rate 0.05).

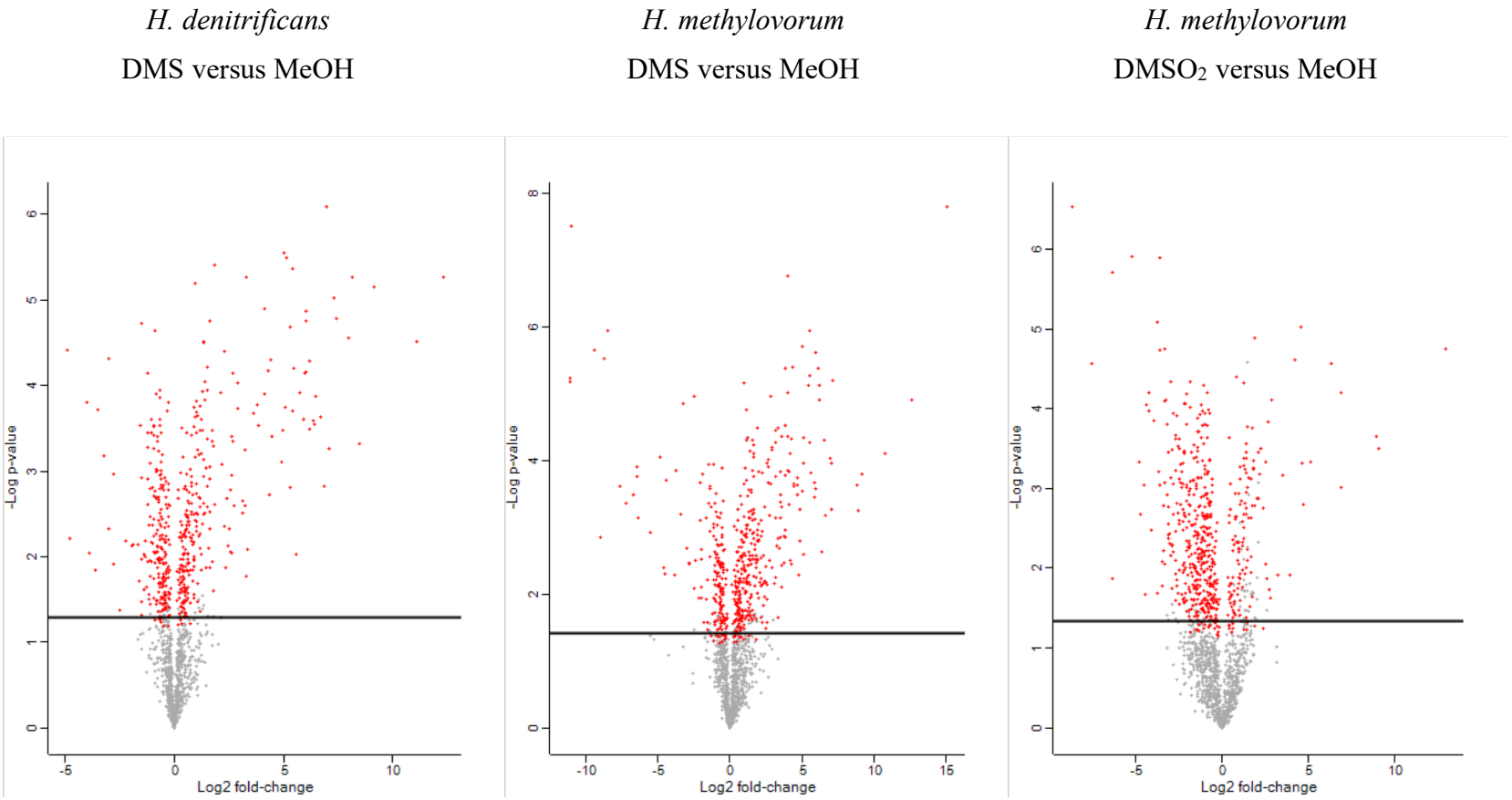


Figure 5.5-5: H. denitrificans *and* H. methylovorum *proteomics volcano plots.* (LEFT) *H. denitrificans* DMS condition (DMS & Na_2SO_4) versus Control condition (MeOH & Na_2SO_4). (MIDDLE) *H. methylovorum* DMS condition (DMS & Na_2SO_4) versus Control condition (MeOH & Na_2SO_4). (RIGHT) *H. methylovorum* $DMSO_2$ condition ($DMSO_2$ & Na_2SO_4) versus Control condition (MeOH & Na_2SO_4). Significantly upregulated proteins with q-value <0.05 are coloured RED.

In the *H. denitrificans* proteomics data 1,266 proteins have been positively identified by the presence of one or more unique peptides, representing ~36% of the total proteome, 122 of which show significant upregulation on the DMS condition versus the MeOH control condition of 2-fold or more (q-value <0.05). Similarly, the 1501 proteins that have been positively identified in the *H. methylovorum* proteomics data, representing 43% of the total (draft) proteome, with 197 proteins being significantly upregulated by 2-fold and 20 more which were highly upregulated by more than 50-fold. As this upregulation is also accompanied by a substantial shift in the number of downregulated proteins compared to the MeOH, we may be able to infer that the change of MeOH control to either DMS or $DMSO_2$ causes substantial changes to the organism's proteome. Note that as imputation has been used to generate protein abundance data for undetected proteins – specifically proteins that were detected in the triplicate samples of at least one condition – especially high fold-changes (>100-fold) are expected for proteins that are abundant under one condition but absent in another.

The proteomics data for each species has been mapped onto the various pathways of MSC utilisation proposed in Section 5.4 and is described below.

5.5.3 Proteomics of MeOH metabolism

H. denitrificans and *H. methylovorum* contain an MxaFI-type MeOH dehydrogenase, which has been positively identified in the proteome of each condition. The comparative proteomics of MeOH metabolism in these species shows no significant differential expression of MxaFI on DMS compared to a MeOH control condition, nor on $DMSO_2$ for *H. methylovorum*. The same is also true for the related MxaJ and MxaG proteins from the MxaFJGI enzymes of each organism. Similarly, the XoxF-type MeOH dehydrogenases of *H. denitrificans* were identified but not differentially expressed on DMS, while the putative XoxF-type enzymes of *H. methylovorum* were not identified in the organism's data sets.

The lack of differential expression seen for either the MxaFI-type or XoxD-type MeOH dehydrogenases on DMS or $DMSO_2$ suggests that, unlike *H. sulfonivorans*, MeOH metabolism in *H. denitrificans* and *H. methylovorum* is likely to be constitutively active.

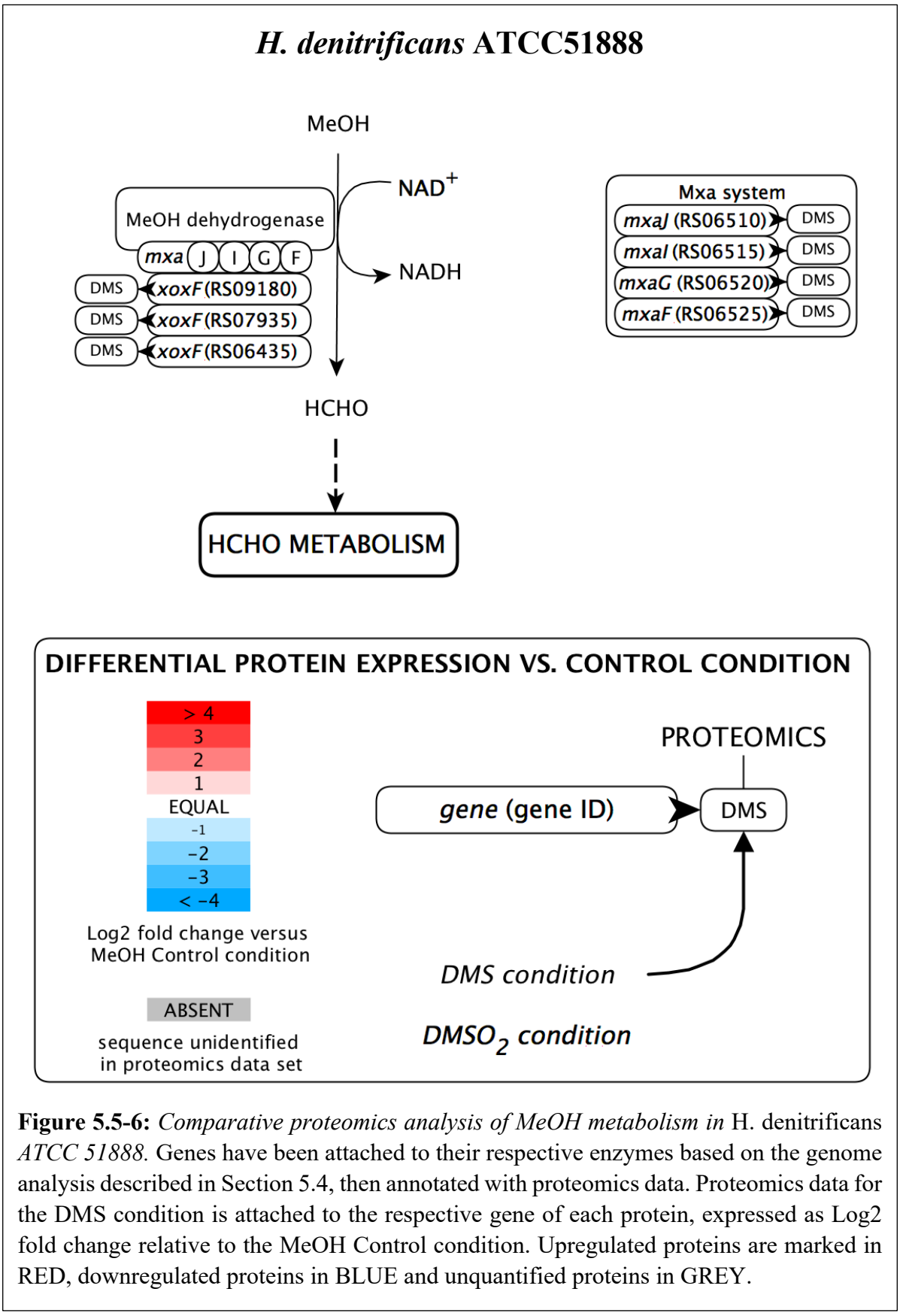


Figure 5.5-6: *Comparative proteomics analysis of MeOH metabolism in* H. denitrificans *ATCC 51888.* Genes have been attached to their respective enzymes based on the genome analysis described in Section 5.4, then annotated with proteomics data. Proteomics data for the DMS condition is attached to the respective gene of each protein, expressed as Log2 fold change relative to the MeOH Control condition. Upregulated proteins are marked in RED, downregulated proteins in BLUE and unquantified proteins in GREY.

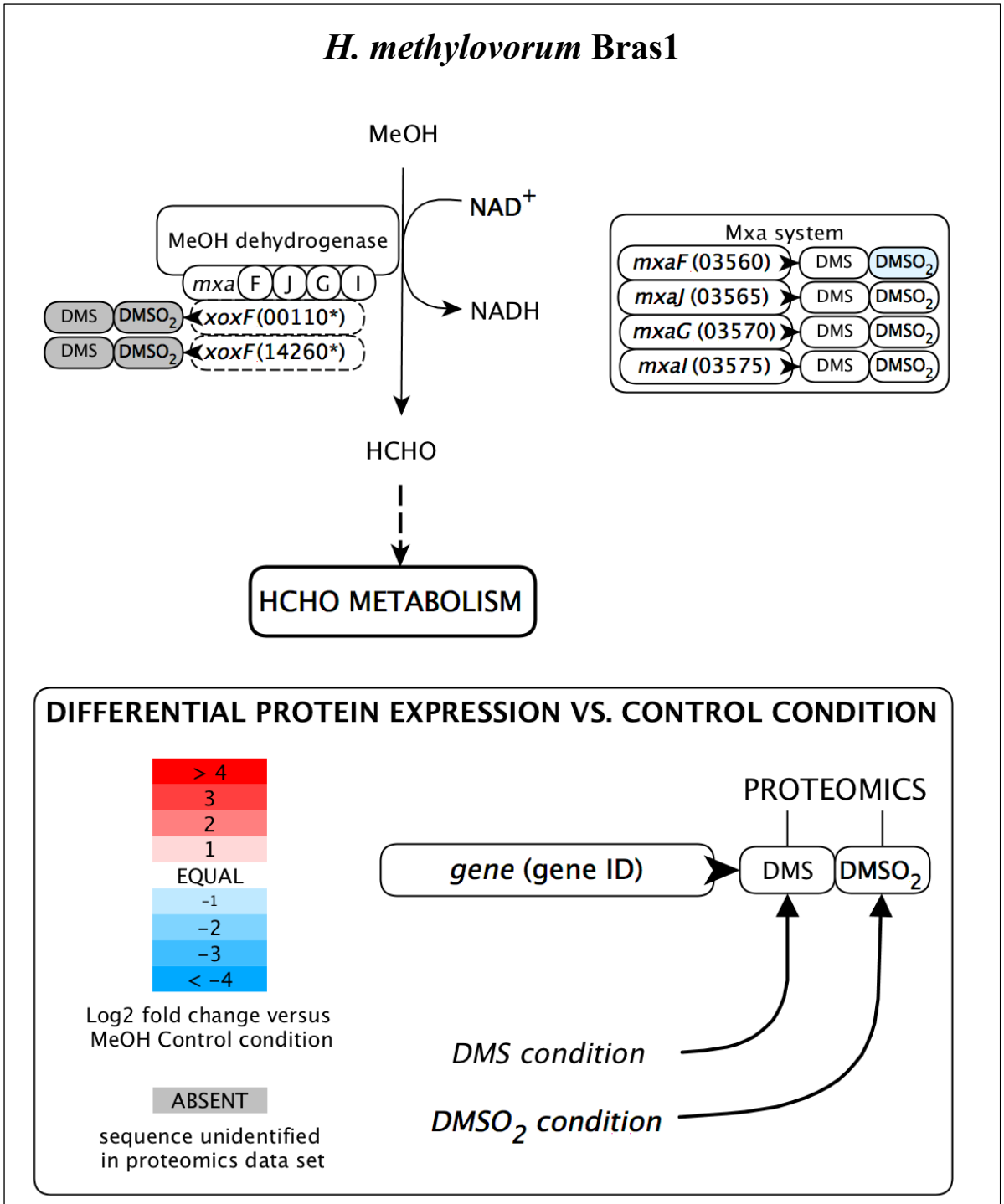


Figure 5.5-7: *Comparative proteomics analysis of MeOH metabolism in* H. methylovorum *Bras1.* Genes have been attached to their respective enzymes based on the genome analysis described in Section 5.4, then annotated with proteomics data. Proteomics data for the DMS and $DMSO_2$ conditions are attached to the respective gene of each protein, expressed as Log2 fold change relative to the MeOH Control condition. Upregulated proteins are marked in RED, downregulated proteins in BLUE and unquantified proteins in GREY. Genes with low confidence are marked with a (*) and dotted outline.

5.5.4 Proteomics of MSC metabolism

The comparative proteomics performed in Chapter 4 shows that when *H. sulfonivorans* was cultivated on a sole carbon source of $DMSO_2$ versus MeOH control, it triggered the induction of two $FMNH_2$-dependent monooxygenases in the organism's proteome: DmoAB-type DMS monooxygenase and a putative SfnFG-like monooxygenase. Both of these enzymes are likely to participate in a methylotrophic pathway of $DMSO_2$ utilisation to generate formaldehyde and either hydrogen sulfide, sulfite or a combination of sulfide and sulfite. However, the precise mechanism of this pathway remains unknown.

MSC metabolism in H. denitrificans

The proteomics data collected for *H. denitrificans* has been mapped onto the metabolic pathway of MSC metabolism constructed in Section 5.4 and displayed below in Figure 5.5-8.

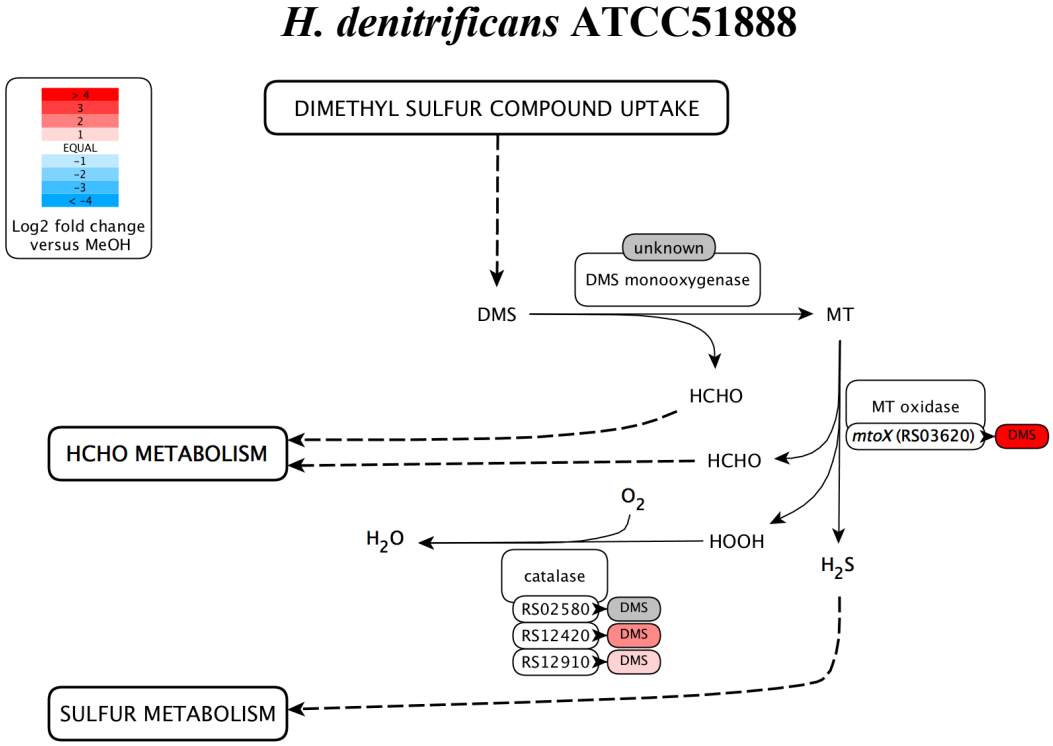


Figure 5.5-8: *Comparative proteomics analysis of MSC metabolism in* H. denitrificans *ATCC 51888.* Genes have been attached to their respective enzymes based on the genome analysis described in Section 5.4, then annotated with proteomics data. Proteomics data for the DMS condition is attached to the respective gene of each protein, expressed as Log2 fold change relative to the MeOH Control condition. Upregulated proteins are marked in RED, downregulated proteins in BLUE and unquantified proteins in GREY.

This previous genome analysis of *H. denitrificans* only identified one known enzyme of MSC metabolism, the putative MtoX-type MT oxidase HDEN_RS03620, which oxidises MT to produce hydrogen sulfide, formaldehyde and hydrogen peroxide.

A search for this enzyme in the proteomics data shows that it is upregulated 570-fold on the DMS condition compared to the MeOH Control condition. Of the three catalases present in the genome, two were upregulated in response to DMS (HDEN_RS12420 and HDEN_RS12910), which is consistent with the expected production of hydrogen peroxide by MT oxidase. Based on the massive upregulation of the DmoAB-type DMS monooxygenase in the *H. sulfonivorans* proteome in response to the growth of the organism on $DMSO_2$, we may expect to see the upregulation of a similar MT producing enzyme when *H. denitrificans* is cultivated on DMS, but the identity of this enzyme is currently unknown.

MSC metabolism in H. methylovorum

The proteomics data collected for *H. methylovorum* has been mapped onto the metabolic pathway of MSC metabolism constructed in Section 5.4 and displayed below in Figure 5.5-9. Genomics suggests that the organism contains four putative enzymes of MSC metabolism: a DmoAB-type DMS monooxygenase, an MtoX-type MT oxidase, an SfnFG-type $DMSO_2$ monooxygenase and an MsmABCD-type MSA monooxygenase. On DMS versus MeOH, the putative-MtoX (DLM45_12340) shows a ~40-fold increase in protein abundance compared to MeOH, while the MsmABCD subunits MsmA (DLM45_05095) and MsmB (DLM45_05100) are upregulated by 35-fold and 14-fold respectively. However there appears to be no significant upregulation of either the MsmC or MsmD subunits, nor the DmoA (DLM45_02930), SfnF (DLM45_02940) or SfnG (DLM45_02945).

On $DMSO_2$ versus MeOH, the putative DmoA of *H. methylovorum* is massively upregulated by ~8000-fold based on the imputated protein abundance values of the MeOH condition, which is consistent with the massive upregulation of DmoA seen previously in the proteomics of $DMSO_2$ metabolism in *H. sulfonivorans*. The abundance of the putative SfnF and SfnG also increases on $DMSO_2$ by ~80-fold and ~12 fold respectively, while the abundance of MtoX (DLM45_12340) is suprisingly downregulated on $DMSO_2$ by a modest ~2 fold. The MsmABCD subunits MsmA (DLM45_05095) and MsmB (DLM45_05100) are upregulated on $DMSO_2$ by ~120-fold and 35-fold respectively, an even greater increase than on DMS. Although the MsmC shows no significant increase on either $DMSO_2$, the MsmD (DLM45_05110) does show an increase of ~9-fold versus MeOH.

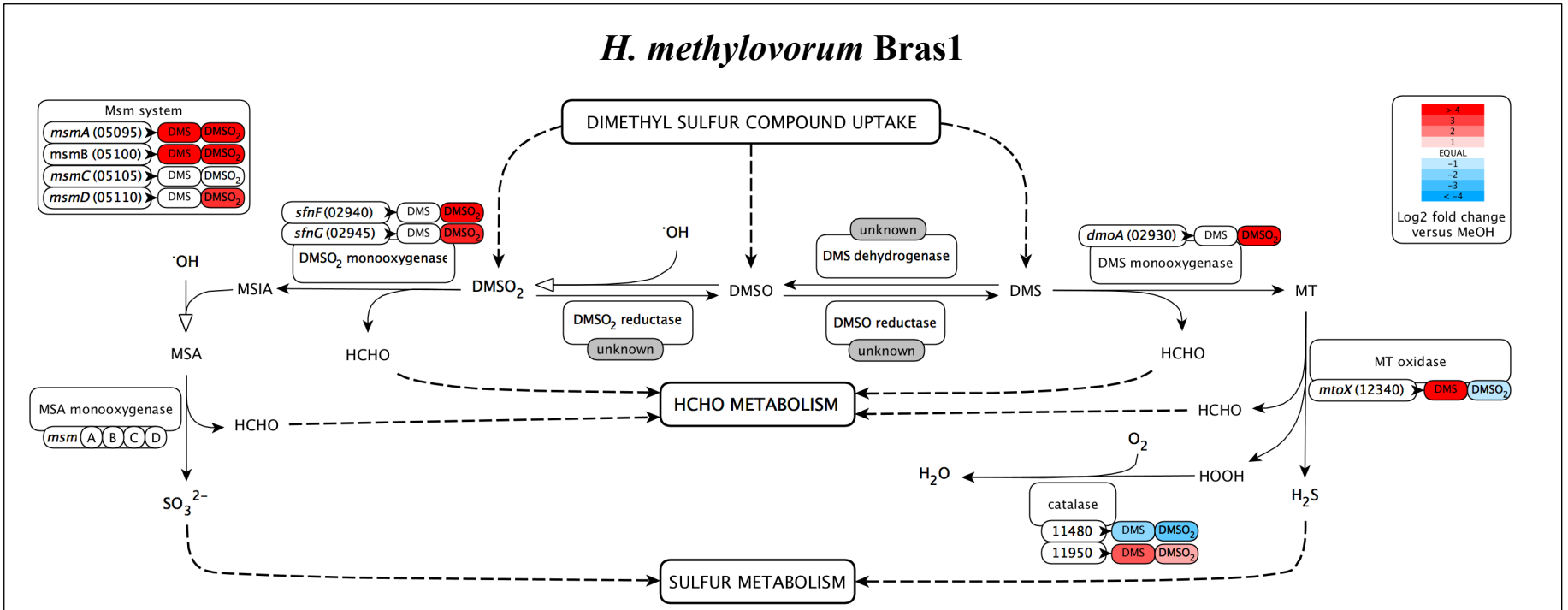


Figure 5.5-9: *Comparative proteomics analysis of MSC metabolism in in* H. methylovorum *Bras1.* Genes have been attached to their respective enzymes based on the genome analysis described in Section 5.4, then annotated with proteomics data. Proteomics data for the DMS and $DMSO_2$ conditions are attached to the respective gene of each protein, expressed as Log2 fold change relative to the MeOH Control condition. Upregulated proteins are marked in RED, downregulated proteins in BLUE and unquantified proteins in GREY.

5.5.6 Proteomics of formaldehyde metabolism

Proteomics of formaldehyde metabolism in H. denitrificans

In *H. denitrificans*, the metabolic pathways and assimilatory or dissimilatory formaldehyde metabolism show no substantial up or downregulation when the organism is cultivated on a sole carbon source of either DMS or MeOH. Each enzyme of the H_4MPT-dependent formaldehyde activation pathway has been successfully identified under both proteomes.

In the H_4F pathway, the FolD-type bifunctional methenyl H_4F cyclohydrolase/ methylene H_4F dehydrogenase HDEN_RS15885 shows a modest ~2-fold downregulation on DMS versus the MeOH control condition. However, as neither of the two putative Fhs-type formate H_4F ligases or other FolD-like enzyme have been identified in the proteomics data set it is unknown to what extent this represents a downregulation of the enzyme or shift in expression to the alternate FolD. Given the close homology between the two Fhs-type enzymes HDEN_RS00520 and HDEN_RS15500 (>90% sequence identity), it is likely that the absence of these enzymes has been caused by a lack of unique peptides rather than genuine lack of expression within the proteome.

Proteomics of formaldehyde metabolism in H. methylovorum

Genome analysis showed that formaldehyde assimilation in *H. methylovorum*, like the other two other model *Hyphomicrobium* species, is likely to involve the generation of 5,10-methylene H_4F via the H_4MPT-dependent formaldehyde activation pathway followed by two duplicates of the H_4F pathway: DLM45_07040/07050 and DLM45_13340/13300. The H_4MPT pathway is present in the organism's proteome when cultivated on a sole carbon source of either $DMSO_2$, DMS or MeOH, with no significant differential protein expression between proteomes. In the H_4F pathway, both Fhs-type formate H_4F ligases and one FolD-like bifunctional methenyl H_4F cyclohydrolase/ methylene H_4F dehydrogenases have been successfully identified.

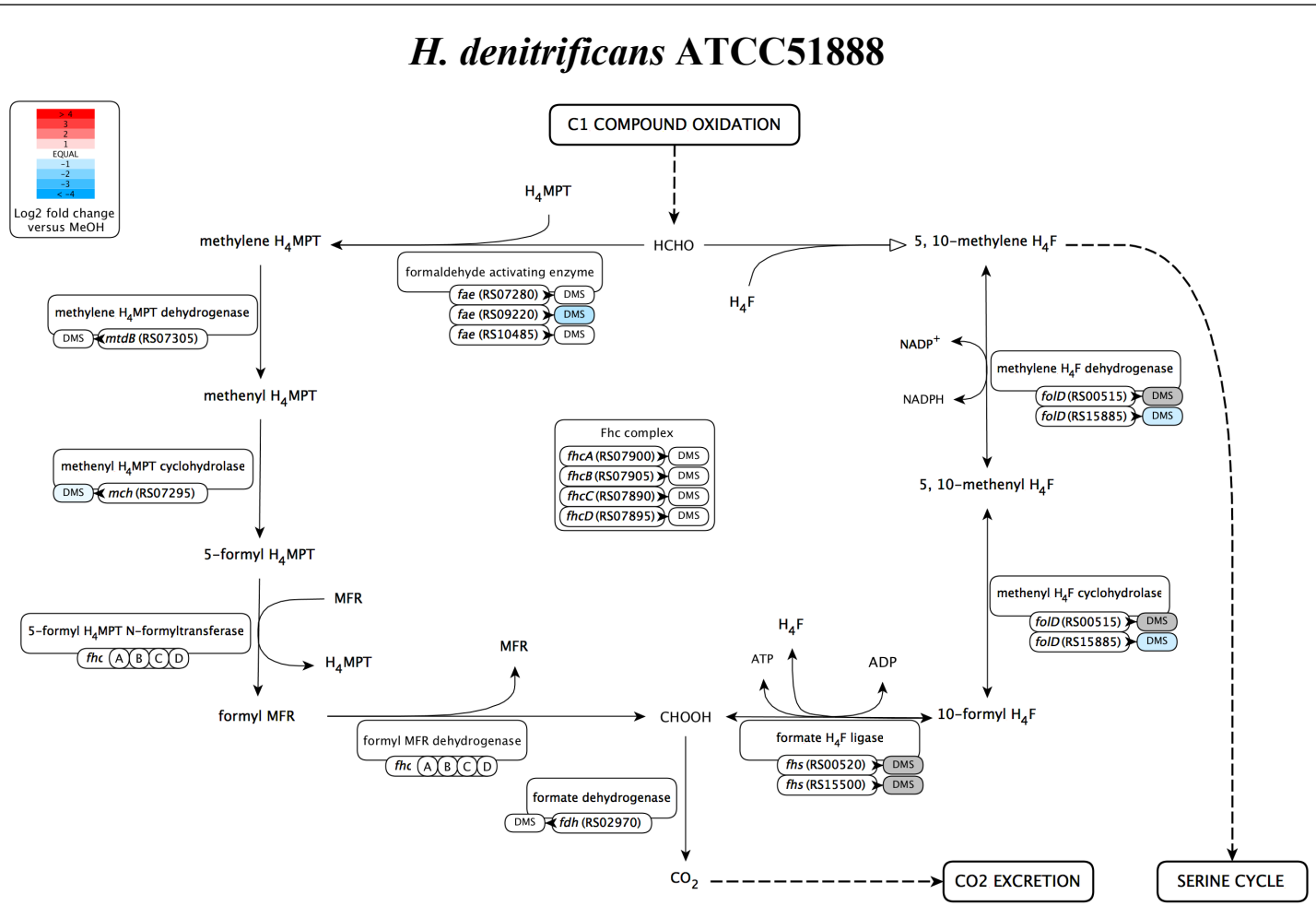


Figure 5.5-10: *Comparative proteomics analysis of formaldehyde metabolism in* H. denitrificans *ATCC 51888.* Genes have been attached to their respective enzymes based on the genome analysis described in section 5.4, then annotated with proteomics data. Proteomics data for the DMS condition is attached to the respective gene of each protein, expressed as Log2 fold change relative to the MeOH Control condition. Upregulated proteins are marked in RED, downregulated proteins in BLUE and unquantified proteins in GREY.

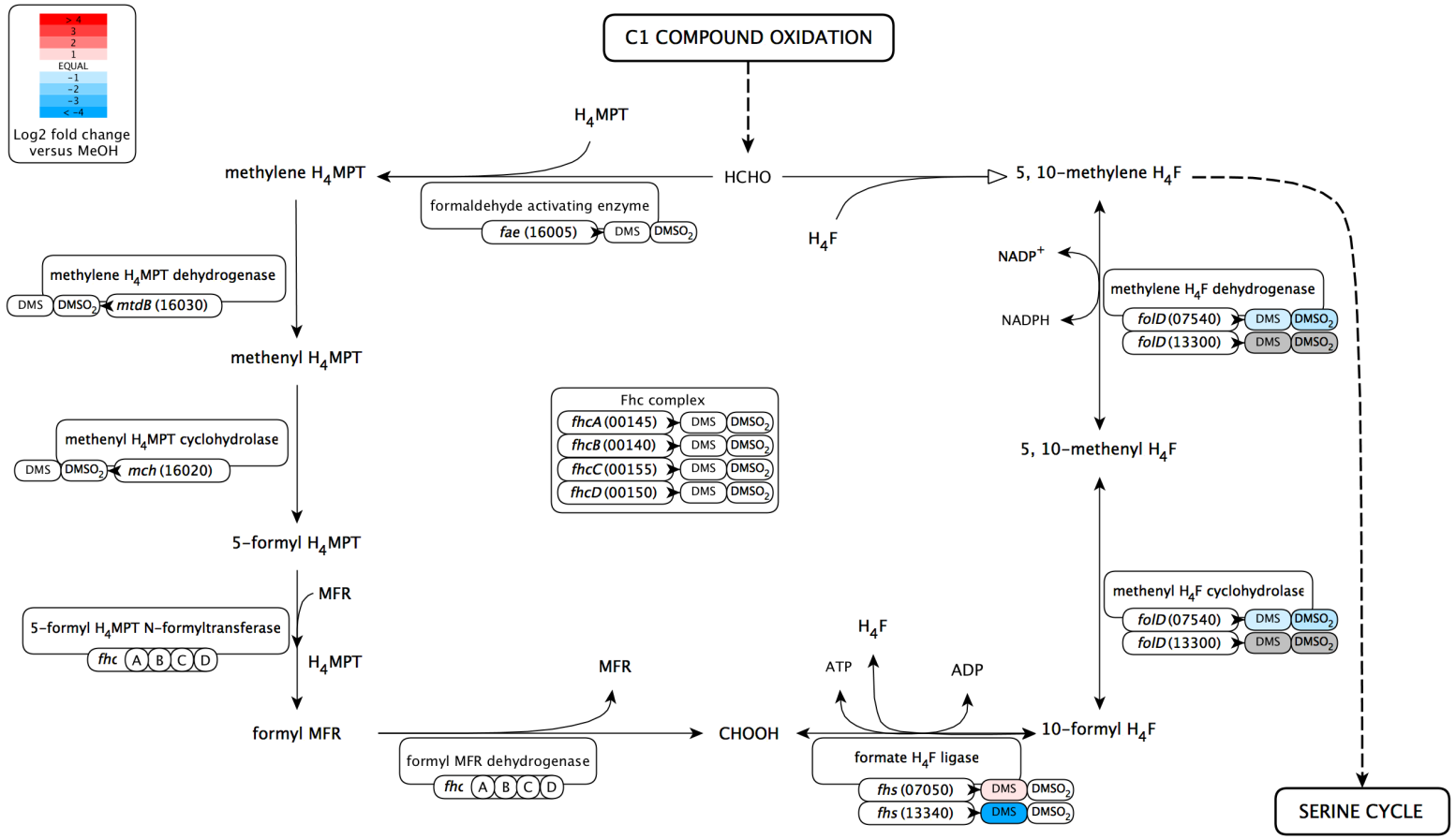


Figure 5.5-11: *Comparative proteomics analysis of formaldehyde metabolism in* H. methylovorum *Bras1.* Genes have been attached to their respective enzymes based on the genome analysis described in section 5.4, then annotated with proteomics data. Proteomics data for the DMS and $DMSO_2$ conditions are attached to the respective gene of each protein, expressed as Log2 fold change relative to the MeOH Control condition. Upregulated proteins are marked in RED, downregulated proteins in BLUE and unquantified proteins in GREY.

The DLM45_07040/07050 pair of Fhs (DLM45_07050) and FolD (DLM45_07040) showed modest differential expression on DMS and $DMSO_2$ versus MeOH (~2-fold change), the other Fhs-type formate H_4F ligase (DLM45_13340) was highly downregulated by ~360-fold on DMS compared to MeOH. This suggests that the second, DLM45_13340/13300 H_4F pathway may either be specifically upregulated in response to MeOH, or else repressed in response to DMS, although it is unknown why this would be the case given the lack of differential expression seen in the other *Hyphomicrobium* H_4F pathways. If this is indeed a DMS-specific downregulation of the pathway then It is unfortunate that mass spectrometry was unable to identify the other H. *methylovorum* FolD, or the absent members of the *H. denitrificans* H_4F pathway described above.

5.5.7 Proteomics of inorganic sulfur metabolism

Inorganic sulfur compound metabolism in H. denitrificans

H. denitrificans does not contain any known enzymes of the sulfite producing $DMSO_2$ oxidation pathway, suggesting that methylotrophic DMS metabolism in this organism will yield hydrogen sulfide but not sulfite. As the sole sulfur source of the MeOH control condition is sulfate, then inorganic sulfur metabolism on this condition is likely to revolve around the intake and reduction of extracellular sulfate to hydrogen sulfide for sulfur assimilation.

Comparing the DMS condition with the MeOH control shows that the putative enzymes of all three steps of the sulfate adenylylation pathway, as well as the CysJ-type sulfite reductase, are highly downregulated on DMS (~4-29-fold). At the same time, the data shows a substantial upregulation of several enzyme systems relating to the assimilation and dissimilation of hydrogen sulfide.

This includes the upregulation on DMS of a CysK-type cysteine synthase (~34-fold), the SqrB, Pdo and Rhd-type enzymes of sulfide oxidation (~90-150-fold), the SoxAXCYZ thiosulfate oxidation system (~10-5000-fold) and the HdrAB1B2C1C2 thiosulfate oxidation system (~20-65-fold). The upregulation of these enzymes also appears to be consistent with the comparative proteomics of DMS metabolism in *H. denitrificans* recently published by Kock and Dahl (2018).

In general, this suggests that the growth of *H. denitrificans* on DMS leads to the generation of excess hydrogen sulfide, downregulating the conversion of sulfate to hydrogen sulfide via the sulfate adenylylation and sulfite reduction pathways, but an upregulation in both sulfide oxidation and thiosulfate oxidation.

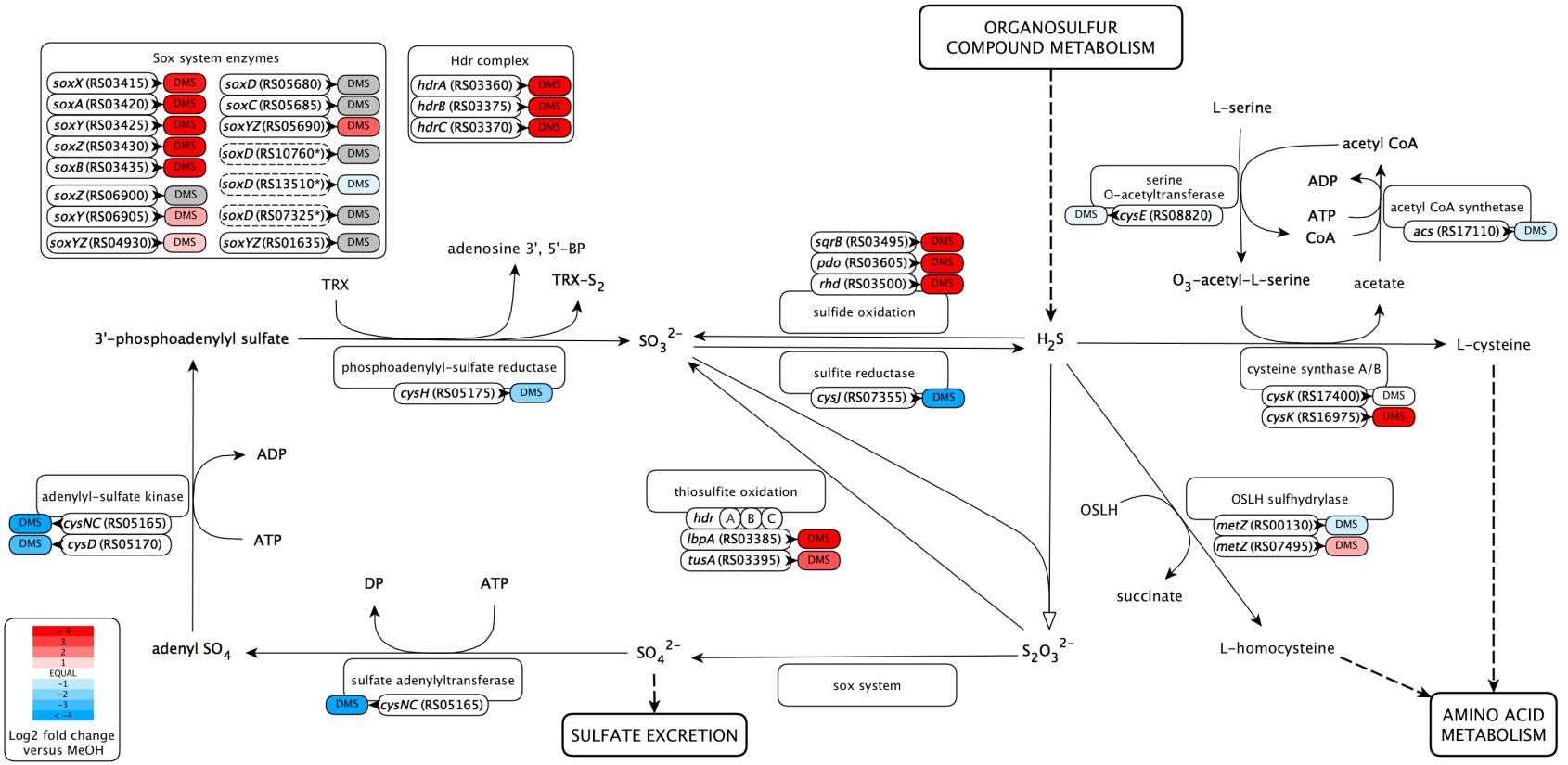


Figure 5.5-12: *Comparative proteomics analysis of inorganic sulfur metabolism in* H. denitrificans *ATCC 51888.* Genes have been attached to their respective enzymes based on the genome analysis described in section 5.4, then annotated with proteomics data. Genes with low confidence are marked with a (*) and dotted outline. Proteomics data for the DMS condition is attached to the respective gene of each protein, expressed as Log2 fold change relative to the MeOH Control condition. Upregulated proteins are marked in RED, downregulated proteins in BLUE and unquantified proteins in GREY. Note the abbreviations OSLH (O-succinyl-L-homoserine) and DP (diphosphate).

Inorganic sulfur compound metabolism in H. methylovorum

As in the two other *Hyphomicrobium* species, sulfur metabolism in *H. methylovorum* is likely to centre around the metabolism of sulfate, sulfite and hydrogen sulfide. Sulfate is the sole sulfur source of the MeOH condition, while sulfite and hydrogen sulfide are the expected products of DMS and $DMSO_2$. Each compound is expected to participate in assimilatory or dissimilatory sulfur metabolism, with assimilatory sulfur metabolism in *H. methylovorum* likely to be mediated by the generation and metabolism of hydrogen sulfide via cysteine and homocysteine synthesis.

On DMS versus MeOH, the proteomics shows a substantial downregulation of the sulfate adenylylation pathway (~15-100-fold) and CysJ-type sulfite reductase (~200-fold), and a more modest downregulation of the CysK-type cysteine synthase A (~9-fold) and one MetZ-type O-succinylhomoserine sulfhydrylase (~5-fold). In contrast, growth on DMS also triggers a massive upregulation of the thiosulfate oxidising SoxXYZABCD clusters DLM45_12160:12195 (average ~1,100-fold) and DLM45_14665:14695 (average ~100-fold), with detected enzymes from each cluster showing statistically significant upregulation. The data also shows that two potential enzymes of sulfide oxidation are also highly upregulated on DMS (Rhd-Sqr and Pdo), that have been identified by virtue of their massive upregulation on DMS and gene synteny with their respective homologues from *H. denitrificans* (discussed below in Section 5.5.8).

On $DMSO_2$ versus MeOH, there is only a moderate downregulation of the sulfate adenylylation pathway (~0-5 fold), CysJ-type sulfite reductase (~3-fold) and CysK-type cysteine synthase A (~5-fold), and MetZ-type O-succinylhomoserine sulfhydrylase (~2-fold). Unlike DMS, $DMSO_2$ only shows a limited upregulation of the SoxXYZABCD clusters DLM45_12160:12195 (average ~4-fold) and DLM45_14665:14695 (average ~1.1-fold), although it does trigger the upregulation of another SoxC (~19-fold) encoded by the *dmoA*/*sfnG* gene cluster. It may be interesting to note that a SoxCD of the *H. sulfonivorans dmoA*/*sfnG2* cluster was also previously found to be upregulated on $DMSO_2$ versus MeOH in Chapter 4, suggesting that this putative SoxCD may actually have a distinct role to that of the other Sox system enzymes. I.e. not necessarily thiosulfate oxidation due to the repression of the other Sox system enzymes in *H. methylovorum*.

Taken together, this suggests that the growth of *H. methylovorum* on DMS causes a shift from assimilatory sulfate metabolism (via sulfate adenylylation and sulfite reduction) to dissimilatory sulfide metabolism (via sulfide oxidation and the Sox system), but that this is not the case when the organism is cultivated on $DMSO_2$.

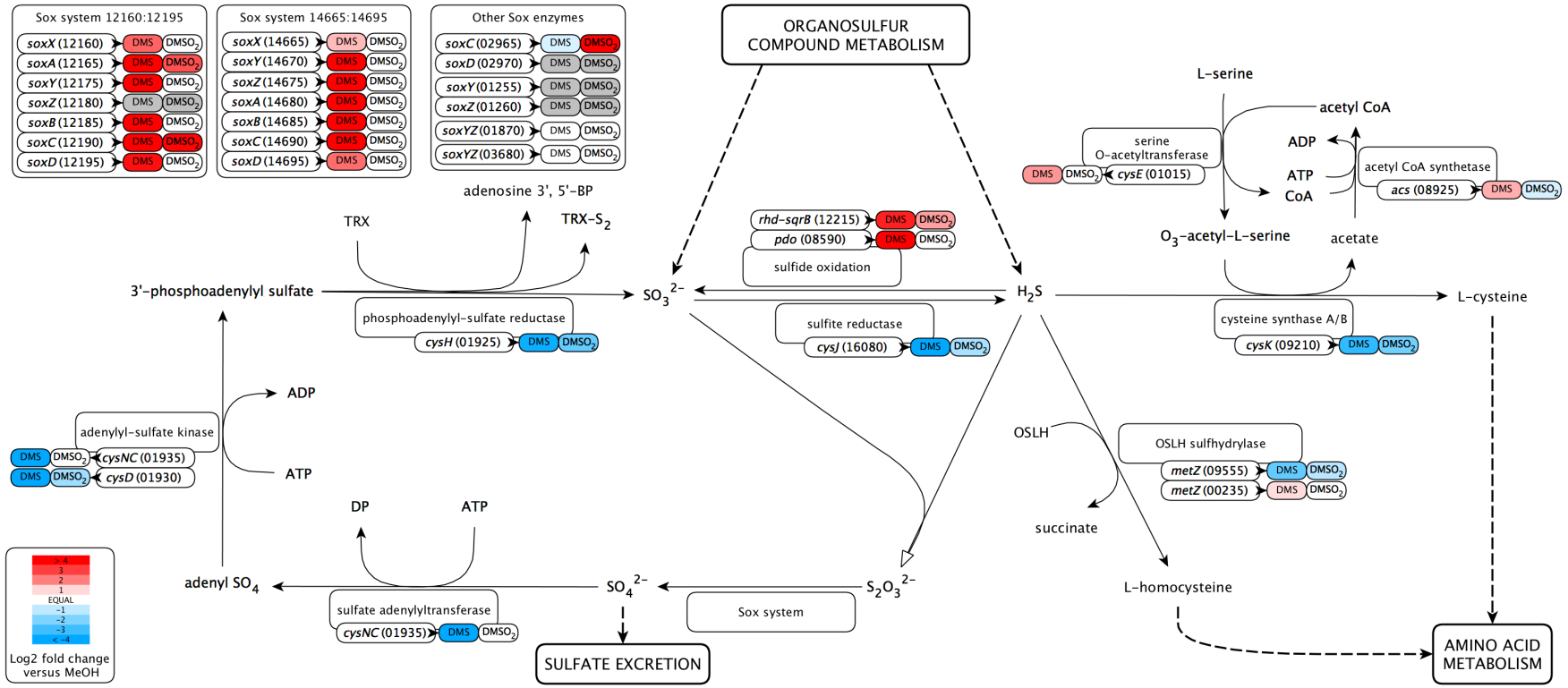


Figure 5.5-13: *Comparative proteomics analysis of inorganic sulfur metabolism in* H. methylovorum *Bras1.* Genes have been attached to their respective enzymes based on the genome analysis described in section 5.4, then annotated with proteomics data. Genes with low confidence are marked with a (*) and dotted outline. Proteomics data for the DMS and $DMSO_2$ conditions are attached to the respective gene of each protein, expressed as Log2 fold change relative to the MeOH Control condition. Upregulated proteins are marked in RED, downregulated proteins in BLUE and unquantified proteins in GREY. Note the abbreviations OSLH (O-succinyl-L-homoserine) and DP (diphosphate).

5.5.8 Other upregulated proteins

DMS-induced proteins in the *H. denitrificans* proteome

In addition to those described in the various metabolic pathways above, numerous other proteins have been found to be highly upregulated in the *H. denitrificans* proteomes when the organism is cultivated on a sole carbon source of DMS compared to the MeOH control (see Table 5.5-1). The vast majority of these proteins appear to be encoded by the same region of the genome spanning ~57.5 kb from the gene HDEN_RS03330 to HDEN_RS03620. Of the 43 proteins predicted to be encoded by this region 34 are upregulated by tenfold or more on DMS versus MeOH.

The upregulated enzymes encoded by this region primarily appear to mediate the oxidation of MT to formaldehyde and the downstream metabolism of the by-products of this reaction, hydrogen sulfide and hydrogen peroxide. This includes enzymes relating to the oxidation of MT to hydrogen sulfide (MtoX), the oxidation of hydrogen sulfide to sulfite (SqrB, Pdo and Rhd), the oxidation of thiosulfate to sulfite (HdrA, HdrB, HdrC, LbpA, SoxY and SoxZ), the oxidation of thiosulfate to sulfate (SoxA, SoxB, SoxX, SoxY and SoxZ) and the detoxification of formaldehyde; although no catalases were found to be highly upregulated in *H. denitrificans* a putative hydrogen peroxide oxidising peroxiredoxin is massively upregulated by >300-fold (HDEN_RS03560).

Additionally, 11 highly upregulated proteins are encoded elsewhere in the genome including two more peroxiredoxins, a CysK-type cysteine synthase and several other proteins of unknown function, due to their poor homology and KEGG orthology.

Table 5.5-1a: *Proteomics heat map of upregulated proteins when* H. denitrificans *is cultivated on a sole carbon source of DMS versus MeOH. Proteins are organised by location within the* H. denitrificans *genome and proteins that have already been described above are highlighted in bold.*

Accession No.	Gene name	NCBI/KEGG annotation	KEGG ID	Strand	Position	fold change versus MeOH Control condition
						DMS
HDEN_RS02650		TonB-dependent receptor		Forward	573515:575794	116.63
HDEN_RS02865		TonB-dependent receptor		Reverse	615813:618176	65.59
HDEN_RS03255		polyphosphate kinase 2		Forward	694546:695640	10.38
HDEN_RS03330	*bioB*	radical SAM protein	K01012	Forward	710755:712056	17.77
HDEN_RS03335		lipoate--protein ligase family protein	K03800	Forward	712080:713141	30.54
HDEN_RS03345		peroxiredoxin		Forward	714007:714384	44.67
HDEN_RS03350	*hdrC*	4Fe-4S dicluster domain-containing protein	K03390	Forward	714410:715153	60.82
HDEN_RS03355	*hdrB*	heterodisulfide reductase subunit B	K03389	Forward	715200:716600	168.90
HDEN_RS03360	***hdrA***	**CoB--CoM heterodisulfide reductase iron-sulfur subunit A family protein**	**K03388**	**Forward**	**716665:717714**	**124.12**
HDEN_RS03370	***hdrC***	**heterodisulfide reductase subunit C**	**K03390**	**Forward**	**718580:719362**	**41.84**
HDEN_RS03375	***hdrB***	**disulfide reductase**	**K03389**	**Forward**	**719411:720355**	**64.61**
HDEN_RS03385	***lbpA***	**glycine cleavage system protein H**	**K02437**	**Forward**	**720866:721306**	**21.58**
HDEN_RS03415	***soxX***	**sulfur oxidation c-type cytochrome SoxX**	**K17223**	**Reverse**	**725184:725852**	**12.54**
HDEN_RS03420	***soxA***	**sulfur oxidation c-type cytochrome SoxA**	**K17222**	**Reverse**	**725830:726663**	**53.70**
HDEN_RS03425	***soxY***	**thiosulfate oxidation carrier protein SoxY**	**K17226**	**Forward**	**726815:727276**	**39.65**
HDEN_RS03430	***soxZ***	**thiosulfate oxidation carrier complex protein SoxZ**	**K17227**	**Forward**	**727314:727643**	**65.72**
HDEN_RS03435	***soxB***	**thiosulfohydrolase SoxB**	**K17224**	**Forward**	**727686:729374**	**5,000.31**
HDEN_RS03495	***sqrB***	**NAD(P)/FAD-dependent oxidoreductase**	**K17218**	**Forward**	**744049:745296**	**102.77**
HDEN_RS03500	***rhd***	**TIGR01244 family phosphatase**	**K17218**	**Forward**	**745472:745921**	**89.03**
HDEN_RS03510	***pdo***	**MBL fold metallo-hydrolase**	**K06897**	**Forward**	**746753:747814**	**157.39**
HDEN_RS03520		hypothetical protein		Reverse	748816:749370	47.27
HDEN_RS03525		hypothetical protein		Reverse	749679:750377	19.57
HDEN_RS03530		DNA-binding response regulator		Reverse	750844:751467	73.24
HDEN_RS03540		acyl-CoA dehydrogenase		Reverse	754021:755187	85.00
HDEN_RS03550		hypothetical protein		Reverse	755589:755867	83.10

colour scale (fold change versus control condition)	<0.1	<1	1	>1	>10	>100	>1000

Table 5.5-1b: *Proteomics heat map of upregulated proteins when* H. denitrificans *is cultivated on a sole carbon source of DMS versus MeOH. Proteins are organised by location within the* H. denitrificans *genome and proteins that have already been described above are highlighted in bold.*

Accession No.	Gene name	NCBI/KEGG annotation	KEGG ID	Strand	Position	fold change versus MeOH Control condition
						DMS
HDEN_RS03555	*azoR*	FMN-dependent NADH-azoreductase	K01118	Reverse	756058:756681	2,143.98
HDEN_RS03560		peroxiredoxin		Forward	756943:757578	359.54
HDEN_RS03565		fatty acid desaturase, alkane 1-monooxygenase	K00496	Forward	758027:759076	255.12
HDEN_RS03570		group 1 truncated hemoglobin	K06886	Forward	759144:759545	72.98
HDEN_RS03575		oxidoreductase FAD/NAD(P)-binding domain protein		Forward	759581:760663	134.88
HDEN_RS03590	*metX*	homoserine O-acetyltransferase/O-succinyltransferase	K00641	Reverse	761641:762687	279.83
HDEN_RS03600		flavin reductase family protein		Forward	763263:763907	31.96
HDEN_RS03605		MBL fold metallo-hydrolase		Forward	764123:765016	39.37
HDEN_RS03610		cytochrome-c peroxidase	K00428	Reverse	765132:766274	22.34
HDEN_RS03615		SCO family protein	K07152	Reverse	766293:766982	43.03
HDEN_RS03620	***mtoX***	**selenium-binding protein**	**K17285**	**Reverse**	**766986:768272**	**570.55**
HDEN_RS04885		TonB-dependent receptor	K02014	Forward	046729:1049209	61.81
HDEN_RS04890		exo-alpha-sialidase		Forward	049213:1050439	17.36
HDEN_RS04895		hypothetical protein		Forward	050436:1050933	35.54
HDEN_RS06415		OsmC family peroxiredoxin		Reverse	326672:1327211	31.45
HDEN_RS11670		peroxiredoxin	K03564	Forward	2425613:2426080	14.43
HDEN_RS11720		hypothetical protein		Forward	2432127:2432606	20.25
HDEN_RS16975	***cysK***	**cysteine synthase A**	**K01738**	**Forward**	**3496108:3497145**	**33.68**

colour scale (fold change versus control condition)	<0.1	<1	1	>1	>10	>100	>1000

DMS-induced proteins in the *H. methylovorum* proteome

As previously seen in *H. denitrificans*, the majority of proteins that are highly upregulated when *H. methylovorum* is cultivated on DMS compared to MeOH are encoded within a single region of the genome. In *H. methylovorum* this is a 54.1 kb gene cluster from DLM45_12090 to DLM45_12350.

The upregulated enzymes encoded by this region primarily appear to mediate the oxidation of MT to formaldehyde and the downstream metabolism of the by-products of this reaction, hydrogen sulfide and hydrogen peroxide. This includes enzymes relating to the oxidation of MT to hydrogen sulfide (MtoX), the oxidation of hydrogen sulfide to sulfite (SqrB, Pdo and Rhd), the oxidation of thiosulfate to sulfite (HdrA, HdrB, HdrC, LbpA, SoxY and SoxZ), the oxidation of thiosulfate to sulfate (SoxA, SoxB, SoxX, SoxY and SoxZ) and the detoxification of hydrogen peroxide, such as peroxiredoxin (HDEN_RS03560). The cluster also contains a putative flavocytochrome *c* (DLM45_12200) which appears to be specifically induced in response to DMS and is located just downstream of the cluster's Sox system. Although no experimentally characterised homologues for this particular flavocytochrome c have been identified in other organisms, an FccAB-type flavocytochrome *c* enzymes has previously been implicated in the oxidation of sulfane sulfur (Lü *et al.*, 2017). A potential role for DLM45_12200 may therefore be sulfide oxidation to sulfur as part of excretory sulfide metabolism.

In addition to this large DMS-sensitive region in the genome are several smaller gene clusters that are also specifically upregulated when the organism is cultivated on DMS versus MeOH, including one which contains the organism's second Sox cluster, but otherwise do not appear to play a direct role in MSC metabolism.

$DMSO_2$-induced proteins in the *H. methylovorum* proteome

Several proteins identified by mass spectrometry in the *H. methylovorum* proteome are either specifically induced in response to the growth of the organism on $DMSO_2$, or on both DMS and $DMSO_2$. The former group includes DmoA and SfnFG which are part of a *dmoA*/*sfnG* gene cluster with striking similarity to the *H. sulfonivorans dmoA*/*sfnG2* cluster found to be induced on $DMSO_2$; in addition to DmoA and SfnFG, the cluster also contains a porin, putative SoxCD, several Rib family enzymes associated with riboflavin biosynthesis and an ABC transporter of unknown function.

A small proportion of the proteins that are highly upregulated on DMS are also highly upregulated on $DMSO_2$, including the MsmABCD-type MSA monooxygenase which is accompanied by an amidohydrolase, and a small gene cluster encoding a putative disulfide bond formation protein, cysteine hydrolase and carboxylesterase. It is interesting this enzyme is upregulated on both $DMSO_2$ and DMS, especially as it has previously been associated with the utilisation of MSA as a carbon source by marine microorganisms, and the phenotyping of *H. methylovorum* performed earlier in this chapter suggests that it is unable to utilise MSA as either a carbon or sulfur source.

Table 5.5-2a: *Proteomics heat map of upregulated proteins when* H. methylovorum *on a sole carbon source of DMS versus MeOH. Proteins are organised by location within the* H. methylovorum *genome and proteins that have already been described above are highlighted in bold.*

Accession No.	Gene name	NCBI/KEGG annotation	KEGG ID	Strand	Position (Contig1)	fold change versus MeOH Control condition	
						$DMSO_2$	DMS
DLM45_00185	-	hypothetical protein		Forward	38698:39183	1.00	18.01
DLM45_00580	-	TonB-dependent siderophore receptor	K02014	Forward	115682:118054	2.70	13.86
DLM45_02200	-	glycosyl hydrolase		Reverse	439438:440661	1.00	16.05
DLM45_02205	-	TonB-dependent receptor	K02014	Reverse	440664:443171	1.00	84.65
DLM45_02930	***dmoA***	**5,10-methylene tetrahydromethanopterin reductase**	**K20938**	**Reverse**	**572266:573705**	**8,014.77**	**1.00**
DLM45_02940	***sfnF***	**hypothetical protein**	**K00299**	**Forward**	**575474:576433**	**82.05**	**1.00**
DLM45_02945	***sfnG***	**dimethyl sulfone monooxygenase SfnG**	**K17228**	**Forward**	**576512:577603**	**11.67**	**1.00**
DLM45_02950	*ribB*	3,4-dihydroxy-2-butanone-4-phosphate synthase	K14652	Forward	577666:578784	546.86	1.00
DLM45_02965	***soxC***	**sulfite dehydrogenase**	**K17225**	**Forward**	**580700:581968**	**19.04**	**0.66**
DLM45_04130	-	TonB-dependent receptor	K02014	Forward	807287:809590	1.00	467.84
DLM45_04415	-	DUF1289 domain-containing protein		Forward	869847:870722	1.00	45.76
DLM45_04755	-	alkylhydroperoxidase		Forward	937651:938013	1.00	10.92
DLM45_05085	-	amidohydrolase	K07045	Reverse	1016346:1017221	118.68	22.32
DLM45_05095	***msmA***	**methanesulfonate monooxygenase**		**Forward**	**1018718:1019977**	**118.05**	**35.19**
DLM45_05100	***msmB***	**hypothetical protein**	**K16969**	**Forward**	**1020060:1020548**	**35.36**	**13.98**
DLM45_05110	***msmD***	**oxidoreductase**	**K15765**	**Forward**	**1020917:1022020**	**9.43**	**1.00**
DLM45_05580	-	haloacid dehalogenase	K17686	Reverse	1107530:1109917	1.00	93.32
DLM45_06450	-	disulfide bond formation protein DsbA		Reverse	1301856:1302689	26.06	11.80
DLM45_06455	-	cysteine hydrolase		Reverse	1302787:1303620	502.87	135.68
DLM45_06555	-	carboxylesterase		Forward	1326487:1327605	5.93	26.15
DLM45_07415	-	hemin receptor	K16087	Reverse	1497624:1499894	1.00	58.89
DLM45_07500	-	methyltransferase type 11		Reverse	1517381:1518190	1.00	25.11

colour scale (fold change versus control condition)	<0.1	<1	1	>1	>10	>100	>1000

Table 5.5-2b: *Proteomics heat map of upregulated proteins when* H. methylovorum *on a sole carbon source of DMS versus MeOH. Proteins are organised by location within the* H. methylovorum *genome and proteins that have already been described above are highlighted in bold.*

Accession No.	Gene name	NCBI/KEGG annotation	KEGG ID	Strand	Position (Contig1)	fold change versus MeOH Control condition	
						$DMSO_2$	DMS
DLM45_08570	-	hypothetical protein		Reverse	1723438:1724238	1.00	61.62
DLM45_08590	-	MBL fold metallo-hydrolase		Forward	1725509:1726402	1.00	142.66
DLM45_09345	-	peptidase M23		Reverse	1896161:1898230	25.45	12.22
DLM45_10675	-	TonB-dependent siderophore receptor	K16090	Reverse	2190541:2192928	1.00	16.11
DLM45_11400	-	sigma-54-dependent Fis family transcriptional regulator		Reverse	2328584:2330092	1.00	11.14
DLM45_11560	-	alpha/beta hydrolase		Reverse	2365400:2366188	1.00	20.81
DLM45_12090	-	DNA-binding response regulator		Reverse	2473009:2473629	1.00	14.11
DLM45_12100	-	acyl-CoA dehydrogenase		Reverse	2476177:2477343	1.00	47.79
DLM45_12115	-	peroxidase		Forward	2478653:2479288	1.00	1,784.92
DLM45_12125	-	hypothetical protein		Reverse	2480260:2480523	1.00	18.72
DLM45_12130	-	hypothetical protein		Reverse	2480569:2481615	1.00	11.18
DLM45_12165	***soxA***	**sulfur oxidation c-type cytochrome SoxA**	**K17222**	**Reverse**	**2485778:2486563**	**1.00**	**18.86**
DLM45_12170	-	MBL fold metallo-hydrolase	K06897	Forward	2486881:2487921	24.09	46.03
DLM45_12175	***soxY***	**thiosulfate oxidation carrier protein SoxY**	**K17226**	**Forward**	**2488407:2488871**	**1.00**	**129.81**
DLM45_12185	***soxB***	**thiosulfohydrolase SoxB**	**K17224**	**Forward**	**2489287:2490981**	**1.00**	**569.94**
DLM45_12190	***soxC***	**sulfite dehydrogenase**	**K17225**	**Forward**	**2490992:2492290**	**15.09**	**6,150.81**
DLM45_12195	***soxD***	**cytochrome C**	**K22622**	**Forward**	**2492229:2492969**	**1.00**	**13.40**
DLM45_12200	-	flavocytochrome C		Forward	2493040:2494320	6.19	33,564.70
DLM45_12210	***pdo***	**MBL fold metallo-hydrolase**		**Forward**	**2494851:2495738**	**1.00**	**19.45**
DLM45_12215	***rhd***	**TIGR01244 family phosphatase**	**K17218**	**Forward**	**2495775:2497442**	**2.77**	**11.18**
DLM45_12230	-	porin		Reverse	2499573:2500838	1.00	11.44
DLM45_12260	-	ABC transporter ATP-binding protein	K01996	Reverse	2507933:2508706	0.08	14.50
DLM45_12265	-	ABC transporter permease	K01999	Reverse	2508786:2510138	0.21	10.25
DLM45_12270	-	branched-chain amino acid ABC transporter permease	K01998	Reverse	2510240:2511319	1.00	12.15
DLM45_12290	-	acyl-CoA dehydrogenase		Reverse	2514668:2515834	1.00	29.47
DLM45_12295	-	hypothetical protein		Reverse	2515923:2516225	1.00	31.90

colour scale (fold change versus control condition)	<0.1	<1	1	>1	>10	>100	>1000

Table 5.5-2c: *Proteomics heat map of upregulated proteins when* H. methylovorum *on a sole carbon source of DMS versus MeOH. Proteins are organised by location within the* H. methylovorum *genome and proteins that have already been described above are highlighted in bold.*

Accession No.	Gene name	NCBI/KEGG annotation	KEGG ID	Strand	Position (Contig1)	fold change versus MeOH Control condition	
						$DMSO_2$	DMS
DLM45_12300	-	hypothetical protein		Reverse	2516237:2516515	0.30	72.62
DLM45_12305	*azoR*	FMN-dependent NADH-azoreductase	K01118	Reverse	2516739:2517368	1.00	122.53
DLM45_12310	-	alkane 1-monooxygenase, fatty acid desaturase	K00496	Forward	2518080:2519129	0.11	25.37
DLM45_12315	-	group 1 truncated hemoglobin	K06886	Forward	2519216:2519602	1.00	61.22
DLM45_12320	-	oxidoreductase		Forward	2519646:2520725	0.10	46.91
DLM45_12330	-	thiosulfate/3-mercaptopyruvate sulfurtransferase	K01011	Reverse	2521513:2522400	0.16	17.08
DLM45_12335	-	LysR family transcriptional regulator		Reverse	2522712:2523677	1.00	14.52
DLM45_12340	***mtoX***	**selenium-binding protein**	**K17285**	**Forward**	**2523917:2525185**	**0.49**	**43.26**
DLM45_12345	-	electron transporter SenC	K07152	Forward	2525217:2525906	1.00	74.20
DLM45_12350	-	cytochrome C peroxidase	K00428	Forward	2525906:2527081	1.00	27.40
DLM45_13120	-	dihydroorotase	K01465	Reverse	2698453:2699787	1.00	10.04
DLM45_13410	-	aldehyde-activating protein		Reverse	2753431:2753991	1.00	14.63
DLM45_14130	-	pseudoazurin		Reverse	2908861:2909316	1.00	11.52
DLM45_14315	-	hypothetical protein		Forward	2954568:2956904	1.00	58.97
DLM45_14610	-	hypothetical protein		Forward	3024366:3024890	1.00	22.25
DLM45_14615	-	hypothetical protein		Forward	3024914:3026386	1.38	16.25
DLM45_14620	-	hypothetical protein		Forward	3026477:3026899	1.00	16.91
DLM45_14625	-	hypothetical protein		Forward	3026941:3027555	1.00	28.28
DLM45_14670	***soxY***	**thiosulfate oxidation carrier protein SoxY**	**K17226**	**Forward**	**3036721:3037176**	**1.00**	**96.63**
DLM45_14675	***soxZ***	**thiosulfate oxidation carrier complex protein SoxZ**	**K17227**	**Forward**	**3037201:3037527**	**1.00**	**468.87**
DLM45_14680	***soxA***	**sulfur oxidation c-type cytochrome SoxA**	**K17222**	**Forward**	**3037585:3038436**	**1.00**	**32.38**
DLM45_14685	***soxB***	**thiosulfohydrolase SoxB**	**K17224**	**Forward**	**3038519:3040210**	**1.00**	**70.32**
DLM45_14690	***soxC***	**sulfite dehydrogenase**	**K17225**	**Forward**	**3040302:3041600**	**1.00**	**33.65**
DLM45_16055	-	pyrroloquinoline quinone biosynthesis protein PqqE	K06139	Reverse	3330164:3331324	1.00	11.87

colour scale (fold change versus control condition)	<0.1	<1	1	>1	>10	>100	>1000

5.5.9 Conclusions

Although the three *Hyphomicrobium* species show major differences in their potential pathways of MSC metabolism, comparative proteomics reveals a common pattern of protein abundance between species when utilising either DMS or $DMSO_2$ as a carbon source. This is to say that orthologous enzymes appear to be upregulated in different organisms when cultivated on DMS or $DMSO_2$, namely when *H. denitrificans* and *H. methylovorum* are cultivated on DMS or *H. methylovorum* and *H. sulfonivorans* are cultivated on $DMSO_2$.

For example, when either *H. denitrificans* or *H. methylovorum* utilises DMS as a sole carbon source the MtoX-type MT oxidase of each organism shows massive upregulation compared to a MeOH control condition. Equally, when either *H. methylovorum* or *H. sulfonivorans* utilises $DMSO_2$ as a sole carbon source each organism shows a massive upregulation of DmoA, SfnF and SfnG-like proteins compared to growth on MeOH. Note that due to the high number of SfnFG-like enzymes in *H. sulfonivorans* it is currently unknown whether this particular enzyme is an ortholog of the SfnFG-like enzyme of *H. methylovorum*, but is an interesting property nonetheless.

This divide between DMS and $DMSO_2$ specific enzymes impacts both the DMS and $DMSO_2$ oxidation pathways, potentially undermining the role of DmoAB-type enzymes and suggesting that an alternate enzyme of DMS oxidation or demethylation may be driving the conversion of DMS to MT in *H. denitrificans* and *H. methylovorum*, while raising further questions as to the identity of the missing MT oxidase-like enzyme in *H. sulfonivorans*. The implications of this observation for the proposed pathways of MSC metabolism in each of these species, and across the *Hyphomicrobium* in general, will be discussed in greater detail in Section 5.7.

5.6 Phylogenetics of MSC metabolism in *Hyphomicrobium* species

The functional genomics of *Hyphomicrobium* species has quite surprisingly demonstrated that *H. denitrificans*, *H. methylovorum* and *H. sulfonivorans* make use of a variety of different enzymes in the catabolism of MSCs to formaldehyde, sulfite and/or hydrogen sulfide, as well as the assimilatory and dissimilatory metabolism of these products. Of particular interest are those enzymes which appear to play major roles in the MSC metabolism of some *Hyphomicrobium* species but are absent from the genomes of others, including the DmoA, MsuD and SfnG-type MSC monooxygenases, MtoX-type MT oxidase and the various components the Sox system. It may therefore be interesting to use phylogenetics to examine the relationships between these enzymes and compare them to the putative enzymes of other bacterial species.

5.6.1 FMNH2-dependent monooxygenases

The phylogenetic analysis of *H. sulfonivorans* S1 $FMNH_2$-dependent monooxygenases in Section 4.4 demonstrated significant differentiation between the putative large subunits of DmoAB-type DMS monooxygenases, MsuD/SsuD-type MSA/alkanesulfonate monooxygenases and SfnFG-type $DMSO_2$ monooxygenases. The genomics of DMS and $DMSO_2$ metabolism performed in this Chapter identified several homologues of these enzymes in *H. denitrificans* ATCC 51888, *H. facile* Bras3, *H. methylovorum* Bras1 and *H.* sp. VS, several of which appear to play an important role in MSC metabolism based on functional genomics.

An alignment was performed for each of the putative $FMNH_2$-dependent monooxygenases from the five *Hyphomicrobium* species, plus the additional monooxygenases sequences from *Hyphomicrobium facile* sp. nov and *Hyphomicrobium* sp. MC1 that were previously found over the course of the phylogenetic analysis of these proteins in Section 4.4. This generated an aligned data set of 21 *Hyphomicrobium* sequences, which were used to construct a phylogenetic tree using bootstrapping (bootstrap 100).

Results

The phylogenetic tree for this analysis showing each of the *Hyphomicrobium* species sequences is displayed in Figure 5.6-1, while an expanded tree incorporating the other bacterial $FMNH_2$-dependent monooxygenase sequences explored in Section 4.4 is shown in

Supplementary Figure S5.6-1. The analysis shows good segregation of $FMNH_2$-dependent monooxygenases into four broad clades containing DmoA-like sequences, SfnG-like sequences, MsuD/SsuD-like sequences and finally an outgroup of sequences with no homology to known enzymes. The only exception is a *H. sulfonivorans* LadA-type monooxygenase previously discussed in Chapter 4, which is most closely related to the sequences of the DmoA clade.

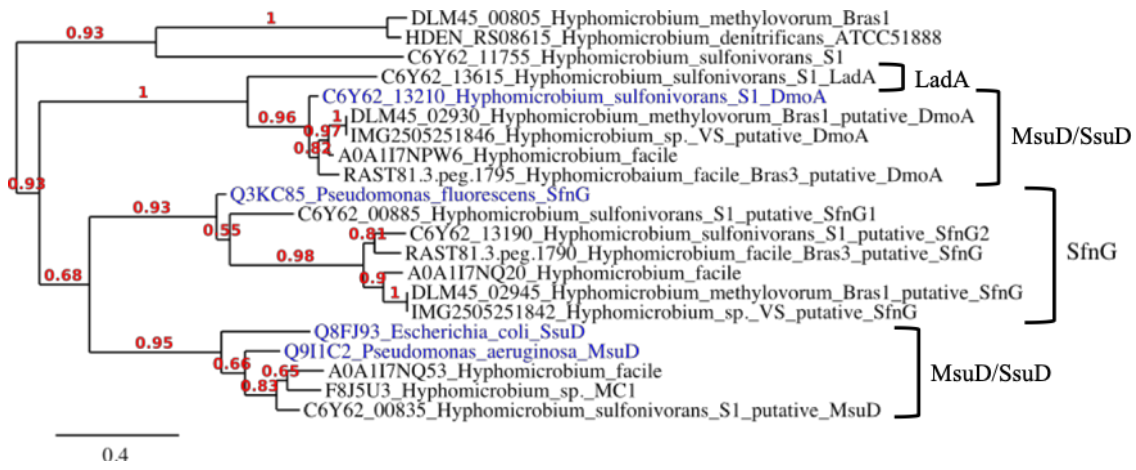


Figure 5.6-1: *Phylogenetic tree of $FMNH_2$-dependent monooxygenase large subunits in* Hyphomicrobium *species.* Experimentally characterised reference sequences are labelled BLUE.

Conclusions

Based on the strong clustering of DmoA, SfnG and MsuD-like *Hyphomicrobium* sequences, we may be able to infer a high chance of functional homology within each clade of $FMNH_2$-dependent monooxygenases. A point of interest is the close homology of the SfnG1-type $DMSO_2$ monooxygenase (C6Y62_00885) with the reference sequence from *P. fluorescens*, which is a minor outlier from the rest of the *Hyphomicrobium* SfnG-like sequences, including SfnG2 (C6Y62_13190). In the genome of *H. sulfonivorans*, SfnG1 is adjacent to a putative SsuABC alkanesulfonate transporter that has yet to be identified in the genomes of other *Hyphomicrobium* species. This suggests that SfnFG1 and SsuABC could be the products of a horizontal gene transfer event between *H. sulfonivorans* S1 and another bacterial species, but without a wider selection of *Hyphomicrobium* reference genomes it cannot be resolved for now.

As none of the outgroup sequences from *H. sulfonivorans* (C6Y62_11755), *H. denitrificans* (HDEN_RS08615) or *H. methylovorum* have been identified by KEGG orthology and none of these had significant differential expression in response to MSCs, it may be safe to assume that these enzymes play no major role in MSC metabolism and are merely the nearest neighbour of $FMNH_2$-dependent MSC monooxygenases in their respective *Hyphomicrobium* species.

5.6.2 MtoX-type methanethiol oxidase

The functional genomics described in Section 5.5 shows that the MtoX-type MT oxidases of *H. denitrificans* and *H. methylovorum* are induced when these organisms are cultivated on a sole carbon source of DMS versus MeOH. The only example of this enzyme to be successfully identified, purified and characterised comes from *H.* sp. VS (Eyice *et al.*, 2017), so it may be interesting to compare this sequence to those of the model *Hyphomicrobium* sequences and MT-like sequences from other bacterial strains.

The procedure for this work follows the method previously described in Section 4.4, in which a protein BLAST search was performed for the MT sequences for the MT homologues taken from each *Hyphomicrobium* sequences against the UNIPROT database. Note that no significant MtoX homologue has been identified in *H. sulfonivorans* S1. A selection of homologous bacterial sequences was then aligned against the *Hyphomicrobium* sequences and used to construct a phylogenetic tree (bootstrap 100).

Results

Phylogenetic analysis of MtoX sequences show goods clustering of the *Hyphomicrobium* sequences into a single clade. The MtoX phylogenetic tree is displayed in Figure 5.6-2.

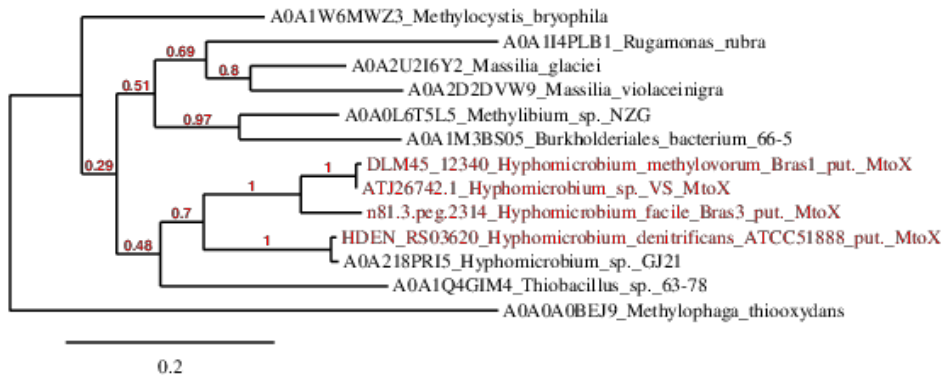


Figure 5.6-2: *Phylogenetic tree of MtoX amino acid sequences from* Hyphomicrobium *species against those from other bacterial species.* Sequences from the four MtoX containing model *Hyphomicrobium* species are labelled RED.

Conclusions

The strong clustering between the various MtoX sequences is as expected of a family of closely related microorganisms. The *H. denitrificans* MtoX is somewhat distinct from the other three *Hyphomicrobium* strains examined in this study, but this is consistent with the results of the preliminary genotyping at the start of this chapter, which suggested that *H. denitrificans* is an outlier from the other three sequences due to its lack of DmoA or SfnG-like sequences.

5.6.3 Msm-type MSA monooxygenase

The preliminary genotyping of *Hyphomicrobium* species described in Section 5.1 showed that, although *H. methylovorum* Bras1 is the only *Hyphomicrobium* species chosen for functional genomics to contain an MsmABCD system, both *H. facile* Bras1 and *H.* sp. VS also contain an Msm-type MSA monooxygenase homologues. A phylogenetic analysis has been performed of the MsmA-type enzymes of these three *Hyphomicrobium* species against other bacterial species; a selection of bacterial MsmA-like sequences have been identified by

performing BLAST searches of the UNIPROT protein database for the putative-MsmA of the three *Hyphomicrobium* species.

A protein BLAST search was performed for the MsmA homologues taken from each of the model *Hyphomicrobium* sequences against the UNIPROT database. Note that no significant MsmA homologues have been identified in either *H. denitrificans* ATCC 51888 or *H. sulfonivorans* S1. A selection of homologous bacterial sequences was then aligned against the *Hyphomicrobium* sequences and used to construct a phylogenetic tree (bootstrap 100).

Results

Phylogenetic analysis of MsmA sequences show good clustering of the *Hyphomicrobium* sequences into a single clade, as previously seen for the MsmA protein sequences. The MsmA phylogenetic tree is displayed in Figure 5.6-3.

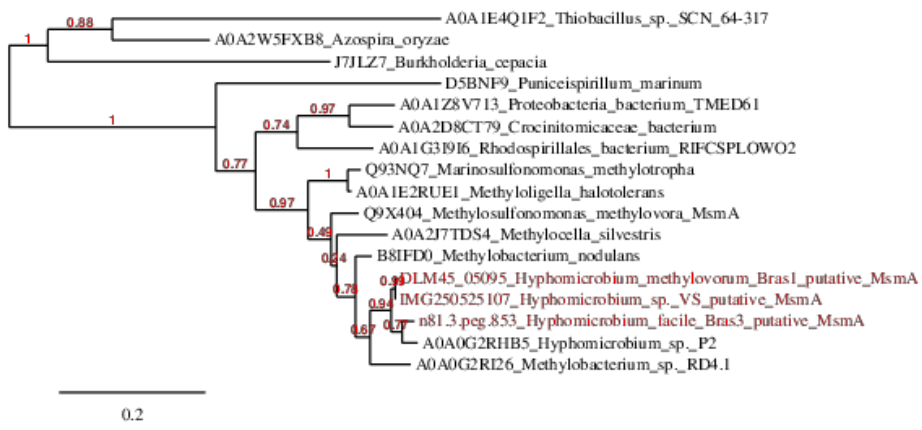


Figure 5.6-3: *Phylogenetic tree of MsmA amino acid sequences from* Hyphomicrobium *species against those from other bacterial species.* Sequences from the three MsmA containing model *Hyphomicrobium* species are labelled RED.

Conclusions

The MsmA sequence hierarchy is consistent with the arrangement of MtoX sequences shown in Figure 5.5-2, with marginal segregation shown between the MsmA sequence of *H. facile* Bras3 and the other two more similar *H. methylovorum* Bras1 and *H.* sp. VS MsmA sequences.

5.6.4 Sox system enzymes

The genotyping of *H. denitrificans* ATCC 51888, *H. methylovorum* Bras1 and *H. sulfonivorans* highlighted a disparity between the Sox system enzymes present in each species. As functional genomics has now linked the induction of several Sox system enzymes with the utilisation of $DMSO_2$ and DMS as a sole carbon source, it may be interesting to examine the relationships between several of these enzymes from the various Sox clusters found in each species.

A protein BLAST search was performed for the SoxA, SoxC and SoxY homologues from *H. denitrificans*, *H. methylovorum* and *H. sulfonivorans* against the UNIPROT database to generate a list of SoxA, SoxC and SoxY amino acid sequences, ranging from poor to strong homologues of the *Hyphomicrobium* sequences. Each of the sequence lists aligned analysed by phylogenetics to construct a phylogenetic tree for SoxA, SoxC and SoxY (bootstrap 100).

SoxA

The Sox system component SoxAX is a heterodimeric cytochrome *c* which transfers thiosulfate onto a cysteine residue of the SoxYZ carrier protein (Fredreich *et al.*, 2005), found in *H. denitrificans* and *H. methylovorum* but not *H. sulfonivorans*. Phylogenetics shows that the *mtoX* gene cluster associated SoxA sequences of *H. denitrificans* (HDEN_RS03420) and *H. methylovorum* (DLM45_12165), while the other *H. methylovorum* SoxA (DLM45_14680) from the organism's additional Sox cluster appears to be associated with a more distinct clade, suggesting it may have been acquired via horizontal gene transfer.

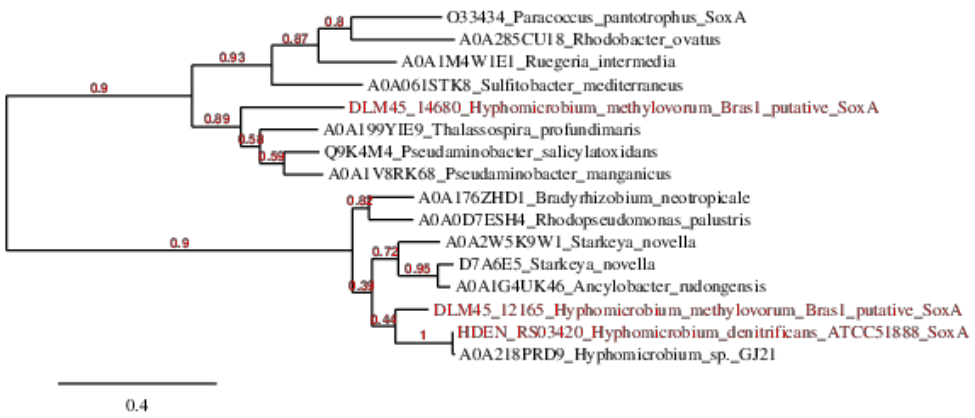


Figure 5.6-4: *Phylogenetic tree of SoxA sequences from* Hyphomicrobium *and other bacterial species. Hyphomicrobium* sequences are labelled in RED.

SoxC

The Sox system component SoxCD is a dehydrogenase that oxidises cysteine-bound S-sulfonate to generate sulfate (Fredreich *et al.*, 2005). The majority of the SoxC *Hyphomicrobium* sequences are separated into distinct clades, with the exception of a single SoxC from *H. sulfonivorans* (C6Y62_13220) and *H. methylovorum* (DLM45_02965). Both of these enzymes are found in the dmoA/sfnG clusters of their respective organisms and appear to be specifically induced in response to the growth of these organisms on $DMSO_2$ as a carbon source.

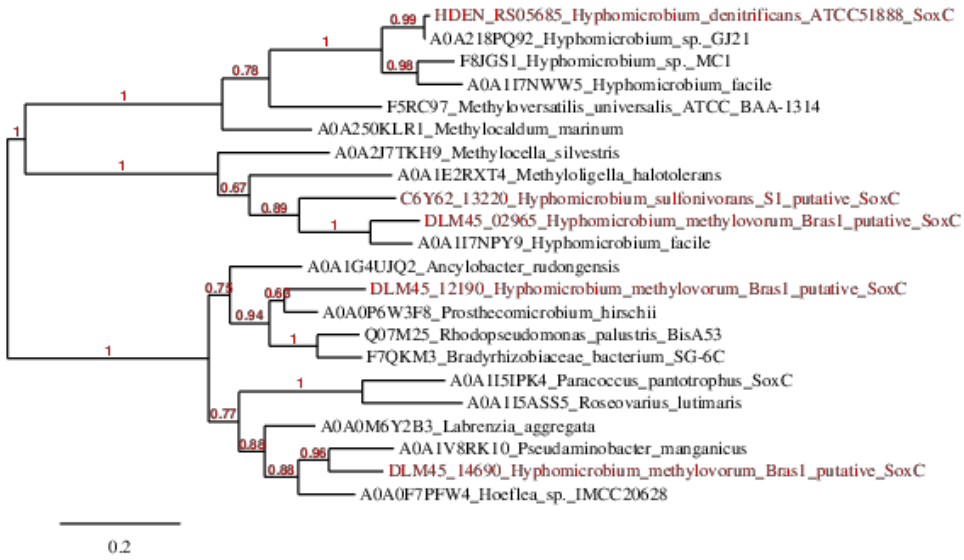


Figure 5.6-5: *Phylogenetic tree of SoxC sequences from* Hyphomicrobium *and other bacterial species. Hyphomicrobium* sequences are labelled in RED.

SoxY

The Sox system component SoxYZ is an inorganic sulfur carrier protein that is associated with Sox and Hdr-mediated thiosulfate oxidation (Fredreich *et al.*, 2005; Koch and Dahl, 2018). The enzyme consists of two subunits, SoxY and SoxZ, found in heterologous pairs in each of the three model *Hyphomicrobium* species' genomes. Phylogenetic analysis suggests that five of these sequences may be divided into two clades of SoxY-like sequences. The first clade contains the SoxY sequences encoded by the *mtoX* gene clusters of *H. denitrificans* (HDEN_RS03425) and *H. methylovorum* (DLM45_12175), while the second contains a conserved SoxY sequence from each species without a clearly defined gene cluster.

The final SoxY sequence (DLM45_14670) is an outlier from the alternative Sox cluster of *H. methylovorum*.

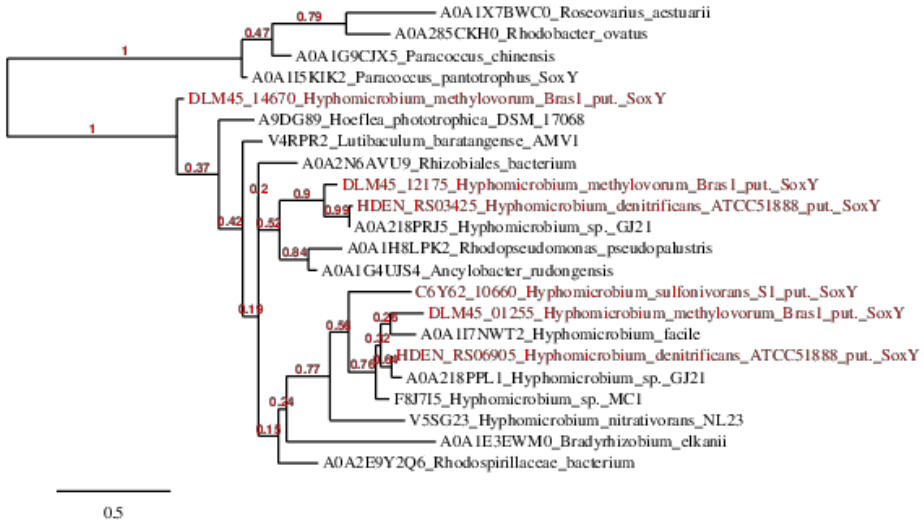


Figure 5.6-6: *Phylogenetic tree of SoxY sequences from* Hyphomicrobium *and other bacterial species. Hyphomicrobium* sequences are labelled in RED.

Conclusions

Phylogenetics appears to support the proposal that the second *H. methylovorum* SoxAXBCDYZ cluster (DLM45_14670:14690) is likely to come from another organism as the result of horizontal gene transfer, based on the consistent outgrouping of the cluster compared to the other *Hyphomicrobium* Sox enzymes. However, the results also suggest that the MSC-induced Sox enzymes found in the *dmoA*/*sfnG* gene cluster and *mtoX* form two distinct but well conserved clades. This suggests that these enzymes are likely to perform the same MSC utilisation-related functions in several different *Hyphomicrobium* species.

5.7 Discussion

The genomics and proteomics of MSC metabolism in *Hyphomicrobium* species described in this chapter demonstrate that *H. denitrificans*, *H. methylovorum* and *H. sulfonivorans* utilise a range of divergent mechanisms for the utilisation of MSCs, as well as the dissimilation of their downstream metabolites. Specifically, it appears that the methylotrophic metabolism of DMS in *H. denitrificans* and *H. methylovorum* involves a different set of MSC degrading enzymes than those used for the methylotrophic metabolism of $DMSO_2$ by *H. methylovorum* and *H. denitrificans*, discussed below.

5.7.1 Mechanisms of methylotrophic DMS metabolism

The methylotrophic metabolism of DMS in *Hyphomicrobium* species has been studied for some time and is thought to be quite well understood, based on previous research performed in *H.* sp. S (De Bont *et al.,* 1981) and *H.* sp. VS (Pol *et al.*, 1994; Eyice *et al.*, 2017). The general consensus is that the mechanism includes an NADH-dependent DMS monooxygenase which oxidises DMS to formaldehyde and MT, and an MT oxidase which oxidises MT to formaldehyde, hydrogen peroxide and hydrogen sulfide. However, the genome analysis and functional genomics of *H. sulfonivorans* performed in Chapter 2 suggests that the pathway is neither ubiquitous nor the sole pathway of MSC metabolism present in *Hyphomicrobium*. A discussion of this pathway in *H. denitrificans* and *H. methylovorum* follows below:

Methylotrophic DMS oxidation pathway

H. methylovorum Bras1 contains a close homologue of both the DmoAB-type DMS monooxygenase of *H. sulfonivorans* (Boden *et al.*, 2011) and an MtoX-type MT oxidase of *H.* sp. VS (Eyice *et al.*, 2017). It was initially believed that the organism would therefore utilise both $DMSO_2$ and DMS via a single DMS oxidation pathway, consisting of this DMS monooxygenase and MT oxidase. However, the functional genomics described in Section 5.5 shows that MtoX is induced when *H. methylovorum* utilises DMS as a sole carbon source and not $DMSO_2$, while the DmoAB-type DMS monooxygenase is upregulated on $DMSO_2$ but not on DMS. This strongly suggests that the utilisation of DMS by *H. methylovorum* is mediated by an alternative enzyme of DMS oxidation instead of DmoAB.

In a similar vein, a search of the *H. denitrificans* genome for a suitable homologue of DmoAB in Section 5.4 yielded no close homologues of the enzyme, or indeed any other known $FMNH_2$-dependent MSC monooxygenase. This also suggests that *H. denitrificans* may be using an alternate enzyme of DMS oxidation in lieu of a DmoAB-type DMS monooxygenase. A summary of this proposed pathway of MSC metabolism is displayed in Figure 5.7-1.

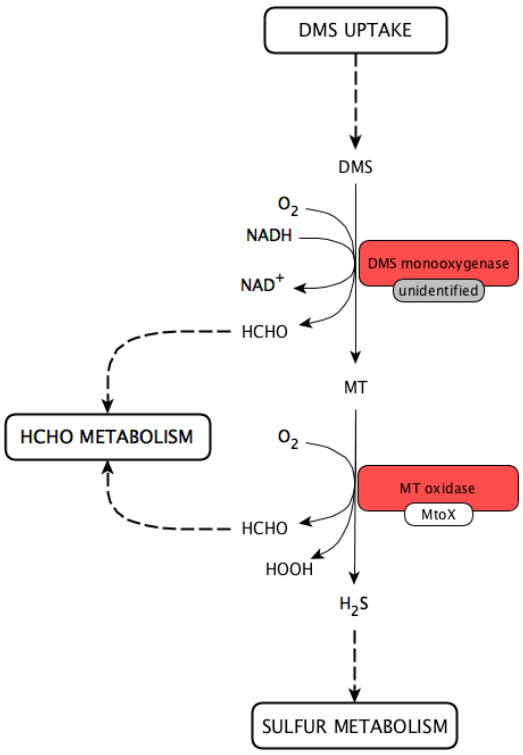


Figure 5.7-1: *Proposed methylotrophic DMS oxidation pathway of* H. methylovorum *Bras1 and* H. denitrificans *ATCC 51888*. Note that as the identity of the DMS monooxygenase remains unknown (GREY).

A potential candidate for this enzyme is the putative AzoR-type FMN-dependent NADH-azoreductase, also suggested by Koch and Dahl (2018), which is encoded within the *mtoX* gene cluster of both *H. denitrificans* and *H. methylovorum*. Like MtoX, AzoR was found in Section 5.4 to be highly upregulated in response to DMS versus MeOH and not on $DMSO_2$, and its use of FMN/NAD cofactors gives it a passing similarity to the known monooxygenases of MSC oxidation. However, the assignment of AzoR as a potential DMS monooxygenase comes with some reservations.

Firstly, experimentally characterised homologues of this enzyme have been found to mediate the cleavage of nitrogen-nitrogen double bonds (azo groups), not thiols (Nakanishi *et al.*, 2001). Given that *H. sulfonivorans* does not contain an MtoX, we may also expect the unknown DMS oxidising enzyme to be absent from the *H. sulfonivorans* genome. As *H. sulfonivorans* contains a putative AzoR based on KEGG orthology and NCBI annotation, encoded ~16 kb downstream of the organisms' *dmoA*/*sfnG2* gene cluster, it could be that AzoR plays an indirect role in MSC metabolism rather than directly oxidising DMS. This enzyme shows no significant differential expression in the proteomics or transcriptomics data between $DMSO_2$ and MeOH/sulfate grown biomass, but as this expression profile is also found for AzoR in *H. methylovorum* it may be inconclusive.

It also appears that some AzoR-type enzymes may function as a quinone reductase, conferring resistance to thiol-specific stress caused by the activity of electrophilic quinones (Liu *et al.*, 2009) and giving them some connection to either hydrogen peroxide detoxification or downstream sulfur metabolism, i.e. sulfide oxidation. Note that thiol stress refers to the capacity of electrophilic quinones to deplete thiol-containing bacterial proteins, not that AzoR itself is involved in thiol reactions (Liebek *et al.*, 2008).

Thiosulfate oxidation is the primary pathway of sulfur dissimilation from DMS

The massive upregulation of MT oxidase seen when *H. denitrificans* and *H. methylovorum* are cultivated on DMS means that hydrogen sulfide is likely to represent a major product of DMS oxidation in these species. Given that bacterial carbon requirements exceed that of sulfur, a likely fate of some portion of this hydrogen sulfide will be excretion from the cell, likely in the form of sulfate via thiosulfate as recently proposed by Koch and Dahl (2018).

The comparative proteomics described in Section 5.5 indicates that several metabolic processes are highly upregulated in response to the growth of *H. denitrificans* and *H. methylovorum* on DMS, mediating the oxidation of hydrogen sulfide to sulfite and the oxidation of thiosulfate, which is consistent with the generation and excretion of excess sulfur from DMS oxidation. The predicted path of hydrogen sulfide oxidation is likely to involve the spontaneous reaction of sulfides and sulfite to generate thiosulfate, followed by Sox-mediated thiosulfate oxidation to generate sulfate for excretion from the cell.

The utilisation of DMS as a sole carbon source also has a secondary effect on sulfur metabolism in which the processes of sulfate adenylation and sulfite reduction appear to be repressed compared to the use of MeOH as sole carbon source instead. As the MeOH control condition utilised sulfate as a sole sulfur source, we can take this repression as evidence that assimilatory sulfate metabolism is reduced on DMS, likely due to the overproduction of hydrogen sulfide by the DMS oxidation pathway; hydrogen sulfide is an intermediate of assimilatory sulfate metabolism via sulfur-containing amino acid synthesis (Sekowska *et al.*, 2000).

Additional mechanisms of dissimilatory sulfur metabolism from DMS

In addition to the common mechanisms discussed above, both *H. denitrificans* and *H. methylovorum* each have an additional enzyme system of sulfur metabolism that is upregulated in response to DMS but specific to each species.

In *H. denitrificans*, this is an electron producing HdrAB1B2C1C2-type heterodisulfide reductase system that is encoded by the same large gene cluster as MtoX and the SoxAXBCDYZ (Koch and Dahl, 2018). This system oxidises thiosulfate to sulfite as part of a sulfite-dependent hydrogen sulfide excretion pathway in which hydrogen sulfide is oxidised to thiosulfate, then further oxidised to sulfate for excretion from the cell. Hdr-mediated thiosulfate oxidation therefore facilitates sulfite recycling during hydrogen sulfide detoxification.

H. methylovorum does not contain a functional heterodisulfide reductase system according to the genome analysis performed in Section 5.4. However, it does contain a secondary SoxAXBCDYZ system that is also induced in response to the growth of the organism on DMS, which is likely to be the result of horizonal gene transfer from another organism. The precise function of the system is currently unknown, although the most likely option is that the system supplements the other SoxAXBCDYZ that can be found encoded by the *mtoX* gene cluster.

5.7.1 Mechanisms of methylotrophic $DMSO_2$ metabolism

It was initially thought that the utilisation of $DMSO_2$ by *H. methylovorum* would follow the methylotrophic DMS oxidation pathway, but the functional genomics of MSC metabolism described in Section 5.5 shows that, unlike on DMS, the putative MtoX-type MT oxidase is not significantly upregulated when the organism in cultivated on $DMSO_2$. This suggests that the organism may actually be using a different mechanism to metabolise $DMSO_2$, as discussed below.

Methylotrophic $DMSO_2$ oxidation pathway

Functional genomics has shown that when *H. methylovorum* Bras1 is cultivated on a sole carbon source of $DMSO_2$, it causes an induction of the putative SfnFG-type $DMSO_2$ monooxygenase, DmoAB-type DMS monooxygenase and a partial induction of the putative MsmABCD-type MSA monooxygenase. This induction of DmoAB and SfnFG on $DMSO_2$ as a carbon source was also found in the functional genomics of *H. sulfonivorans* in Chapter 4,

where it was hypothesised that this SfnFG may either be part of a $DMSO_2$ oxidation pathway and/or may catalyse another MSC oxidation or reduction reaction in methylotrophic $DMSO_2$ metabolism via the DMS oxidation pathway.

However, we have already seen that MSC metabolism via the DMS oxidation pathway causes a major change in the downstream sulfur metabolism of *H. methylovorum* compared to the non-MSC control substrate MeOH; the organism's SoxAXBCDYZ and several enzymes of sulfide oxidation are highly upregulated in response to the excess production of hydrogen sulfide. In contrast, neither of these systems show substantial upregulation when *H. methylovorum* is cultivated on $DMSO_2$, suggesting that hydrogen sulfide is unlikely to be a substantial product of methylotrophic $DMSO_2$ metabolism.

This means that another, non-sulfide producing mechanism must be used by *H. methylovorum* to metabolise $DMSO_2$ as a sole carbon source. Given that a putative $DMSO_2$ monooxygenase and putative MSA monooxygenase are highly expressed when the organism is cultivated on $DMSO_2$, it suggests that a $DMSO_2$ oxidation pathway is far more likely to be the mediator of methylotrophic $DMSO_2$ metabolism using MSA as an intermediate; hence the expected product of a $DMSO_2$ oxidation pathway is sulfite rather than sulfide.

If this is true, then it raises the question of why the DMS monooxygenase subunit DmoA would be massively upregulated in response to $DMSO_2$ if its utilisation does not occur via DMS oxidation? This suggests that either this DmoA plays a different role in *H. sulfonivorans* than *H. methylovorum* or that the enzyme may actually have a different function *in-vivo* than previously found *in-vitro* by Boden *et al.* (2011). The implications of this for the MSC metabolism in *H. sulfonivorans* will be discussed in Chapter 6, but for *H.*

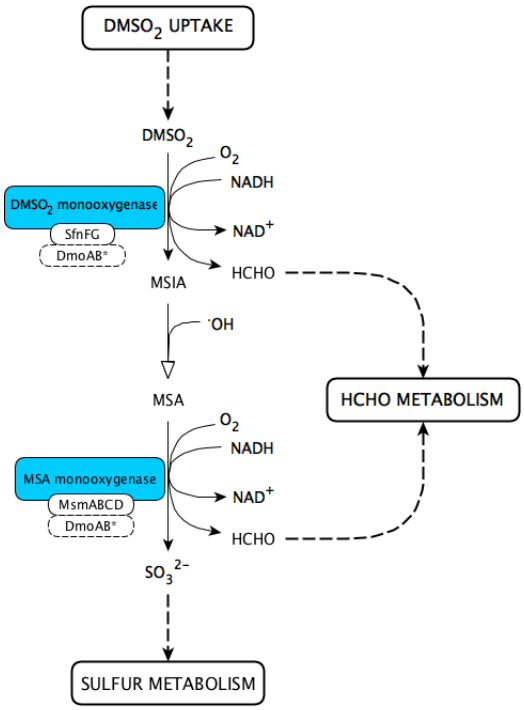


Figure 5.7-2: *Proposed methylotrophic $DMSO_2$ oxidation pathway of* H. methylovorum *Bras1*. Spontaneous reactions are denoted by a white arrowhead. Note that as the function of DmoAB in this pathway remains uncertain, it has tentatively been assigned to the roles of $DMSO_2$ monooxygenase and MSA monooxygenase (*).

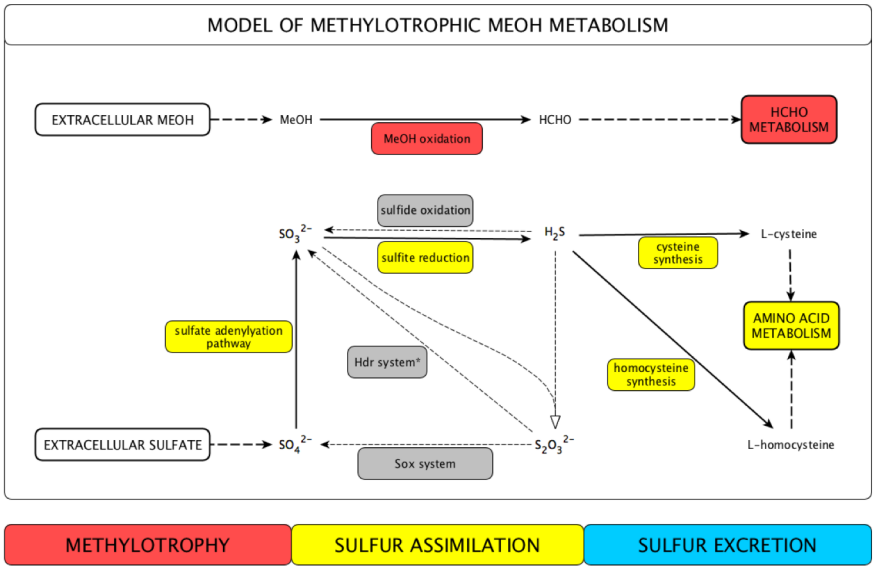
methylovorum then it may suggest that DmoAB mediates either the $DMSO_2$ or MSA oxidation steps of the pathway in conjunction with SfnFG or MsmABCD respectively. A summary of this potential pathway is shown above in Figure 5.7-2.

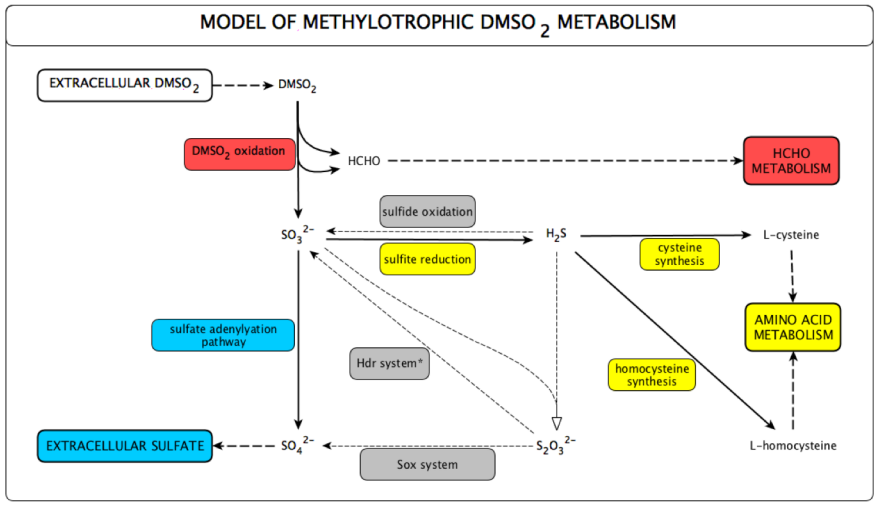
Sulfur adenylylation pathway, an alternate pathway of sulfur dissimilation from $DMSO_2$

As described above, the methylotrophic metabolism of $DMSO_2$ via a $DMSO_2$ oxidation pathway is expected to produce an excess of sulfite analogous to the excess of sulfide produced by the DMS oxidation pathway. However, as the sulfite produced by this pathway does not appear to be utilised via the production and oxidation of thiosulfate, another mechanism is required to generate sulfate for the excretion of excess inorganic sulfur.

In their study of *H. sulfonivorans* S1, Borodina *et al.* (2002) proposed that sulfite oxidation to sulfate may be mediated by the sulfate adenylation system when they were unable to find any evidence of a sulfite dehydrogenase in the organism. As *H. methylovorum* also lacks a sulfite dehydrogenase and neither of the complete SoxAXBCDYZ systems found in the organism's genome are sensitive to $DMSO_2$, it is possible that the sulfate adenylylation system may mediate sulfite oxidation when *H. methylovorum* is cultivated on $DMSO_2$ as a sole carbon source. This would explain why the sulfate adenylylation pathway is highly repressed when *H. methylovorum* is cultivated on DMS versus MeOH but only modestly repressed on $DMSO_2$; the role of the sulfate adenylylation pathway on $DMSO_2$ is to oxidise sulfite to sulfate for excretion from the cell.

The SoxC of a SoxCD encoded by the *dmoA*/*sfnG* cluster is highly upregulated on $DMSO_2$ compared to growth on MeOH or DMS. In the SoxAXBCDYZ system, SoxCD oxidises an S-sulfonate group attached to the sulfur carrier protein SoxYZ to generate sulfate. Phylogenetics found the *dmoA*/*sfnG* cluster SoxC enzymes of both *H. sulfonivorans* and *H. methylovorum* to be distinct from the rest of the Sox system enzymes. Therefore, it is possible this SoxCD may either act as a sulfite dehydrogenase or that the enzyme may cleave sulfate from another, SoxAXB-independent sulfur carrier protein.





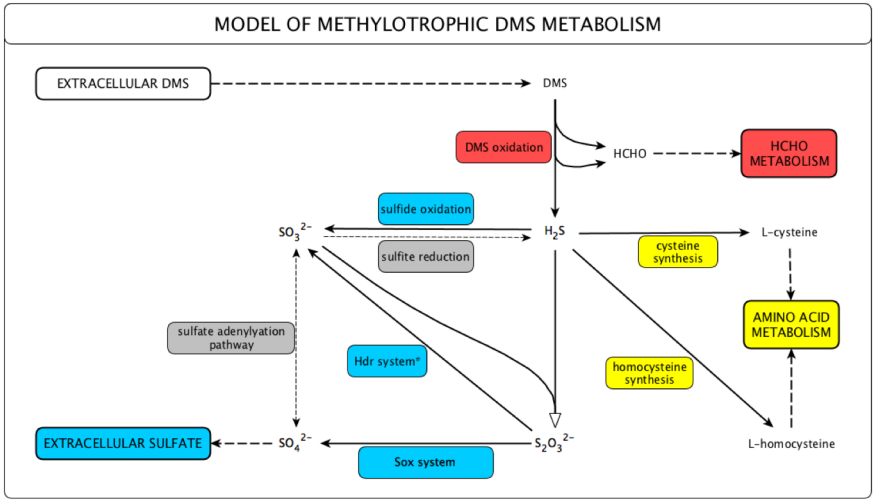


Figure 5.7-3: *Simplified models of inorganic sulfur compound metabolism in* Hyphomicrobium *species under replete sulfate.* (TOP LEFT) Methylotrophic MeOH metabolism as proposed for *H. denitrificans* and *H. methylovorum*. (TOP RIGHT) Methylotrophic $DMSO_2$ metabolism as proposed for *H. methylovorum*. (BOTTOM RIGHT) Methylotrophic DMS metabolism proposed for both *H. denitrificans* and *H. methylovorum*. Active metabolic processes are coloured by function: methylotrophy (RED), sulfur assimilation (YELLOW) and sulfur excretion (BLUE), absent or repressed processes are coloured GREY. Note that *H. methylovorum* does not contain a Hdr system (*) of thiosulfate oxidation, but it has been included on the basis of potential relevance to other *Hyphomicrobium* species.

5.7.3 Conclusions

Returning to the phenotyping of *Hyphomicrobium* species described earlier in this chapter, *Hyphomicrobium* species can make use of a broad range of different MSCs as either a sole carbon or sulfur source. Beginning by drawing a connection between the presence of DmoA and MtoX with either $DMSO_2$ or DMS utilisation, the work described in this Chapter has led to the proposal of two distinct MSC oxidation pathways for future study: the DMS oxidation pathway and $DMSO_2$ monooxygenase pathway.

However, even if these pathways are accurate and an alternative DMS monooxygenase can be identified, our understanding of MSC metabolism in the *Hyphomicrobium* is far from complete. *H. methylovorum* Bras1 and *H.* sp. VS have both been shown to utilise DMSO as a carbon and sulfur source, but it remains unknown whether methylotrophic DMSO metabolism occurs via a DMS or $DMSO_2$ oxidation pathway. Furthermore, the capacity for *H. facile* Bras3 to utilise MSA as a sole carbon source may also be an interesting area of inquiry as the organism contains no strong homologues for either an MsmEFGH or SsuABC-type MSA transport system – potentially suggesting that the organism instead makes use of an alternate and/or novel MSA transporter.

Chapter 6:

Discussion

6. Discussion

6.1 New mechanisms of MSC metabolism proposed for *Hyphomicrobium* species

At the beginning of this project, the metabolic processes mediating methylotrophic MSC utilisation in *Hyphomicrobium* species were thought to be quite well understood. Previous research into *Hyphomicrobium* species such as *H.* sp. S (De Bont *et al.*, 1981), *H. sulfonivorans* S1 (Borodina *et al.*, 2000; Borodina *et al.*, 2002; Boden *et al.*, 2011) and *H.* sp. VS (Pol *et al.*, 1994; Eyice *et al.*, 2017) had generated a plausible model for the metabolism of MSCs in which $DMSO_2$ was reduced to DMS, then fed into a DMS oxidation pathway to produce formaldehyde for methylotrophic carbon assimilation. However, evidence collected over the course of this project now suggests that MSC metabolism in *Hyphomicrobium* is more diverse than first thought.

The functional genomics of $DMSO_2$ metabolism in *H. sulfonivorans* provides compelling evidence for the presence of two distinct mechanisms of MSC utilisation in the organism, (i) a methylotrophic pathway for the utilisation of $DMSO_2$ as a carbon source that is dependent on the activity of DmoAB and (ii) an alternative non-methylotrophic $DMSO_2$ utilisation pathway for the assimilation of MSCs as a sulfur source. Candidate enzymes have been found for the enzymes of both pathways and have been traced back to two gene clusters in the *H. sulfonivorans* genome.

As discussed at the end of Chapter 5, the work performed in this project also suggests that certain *Hyphomicrobium* species can use two distinct mechanisms for methylotrophic MSC metabolism, (i) a $DMSO_2$-induced $DMSO_2$ oxidation pathway which produces formaldehyde and sulfite, and (ii) a DMS-induced DMS oxidation pathway involving MtoX which produces formaldehyde and hydrogen sulfide. Furthermore, functional genomics suggests that the inorganic sulfur oxidation pathways driving the dissimilatory sulfur metabolism of MSCs depends on which of the MSC degradation pathways a strain uses to degrade MSCs (see Figure 6.1-1).

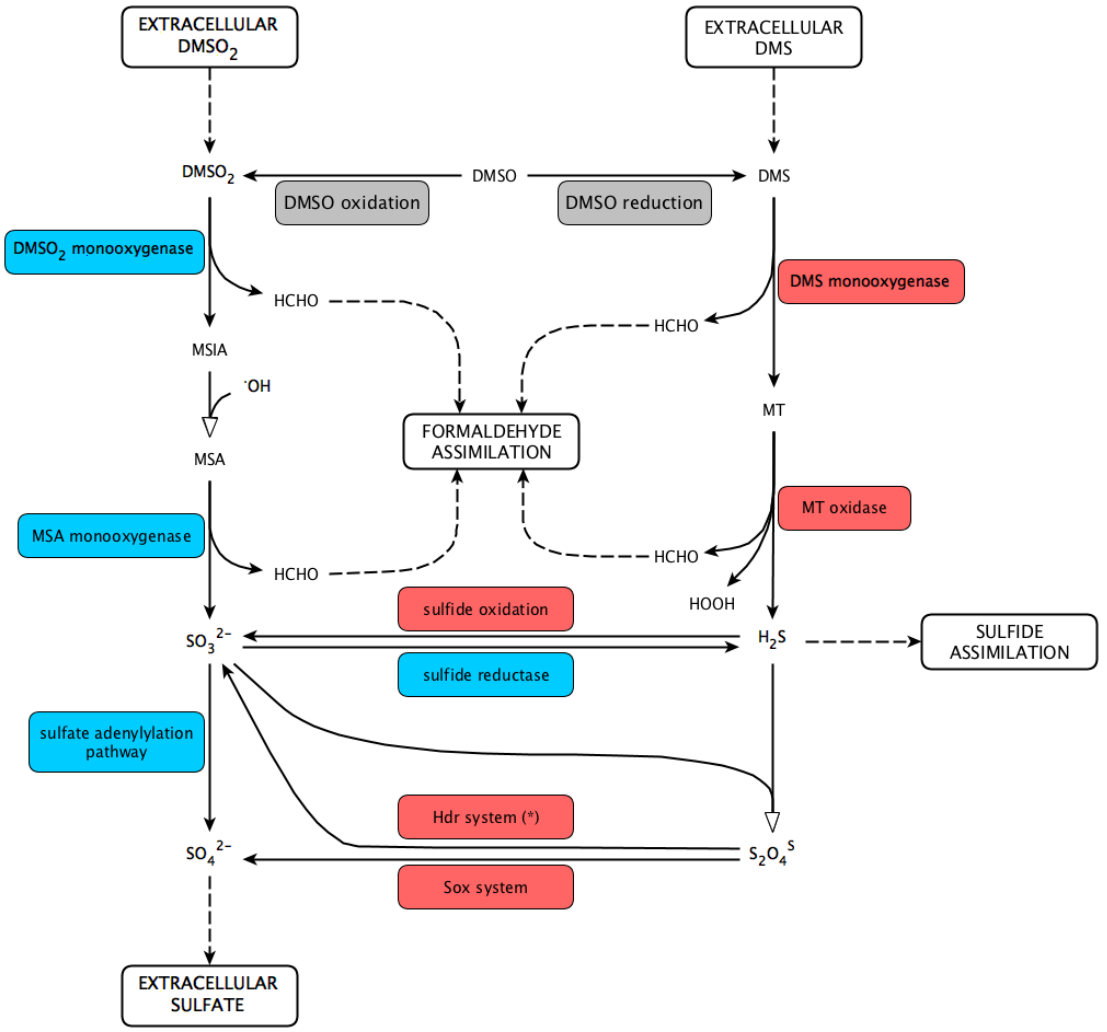


Figure 6.1-1: *Simplified scheme of methylotrophic DMS and DMSO metabolism proposed in* Hyphomicrobium *species*. (RED) Methylotrophic DMS metabolism via a DMS oxidation pathway, in which DMS is oxidised to hydrogen sulfide and formaldehyde. Formaldehyde is assimilated as a carbon source via the serine cycle. Hydrogen sulfide is either assimilated as a sulfur source via amino acid synthesis, or converted to sulfate via thiosulfate oxidation for excretion from the cell. (BLUE) Methylotrophic $DMSO_2$ metabolism via a $DMSO_2$ oxidation pathway, in which $DMSO_2$ is oxidised to sulfite and formaldehyde. Formaldehyde is assimilated as a carbon source via the serine cycle. Sulfite is either reduced to hydrogen sulfide for assimilation as a sulfur source via amino acid synthesis, or converted to sulfate via the reverse sulfate adenylylation pathway for excretion from the cell. (GREY) Methylotrophic DMSO metabolism is likely to occur by either the reduction of DMSO to DMS or the oxidation of DMSO to $DMSO_2$, entering into the DMS or $DMSO_2$ oxidation pathways respectively. Note that the Hdr system (*) of thiosulfate oxidation has so far only been found in *H. denitrificans* ATCC 51888.

6.2 Methylotrophic $DMSO_2$ utilisation pathway of *H. sulfonivorans* S1

One of the main goals of this project has been to identify the unknown enzymes of MSC metabolism in *H. sulfonivorans* S1. To briefly reiterate our understanding of MSC metabolism in *H. sulfonivorans* at the outset of this project, this process was thought to begin with the conversion of $DMSO_2$ to DMS by a pathway of sequential $DMSO_2$ and DMSO reduction (Borodina *et al.*, 2000). This DMS would then be oxidised to MT and formaldehyde by a DMS monooxygenase, and the MT oxidised to formaldehyde, hydrogen peroxide and hydrogen sulfide by some form of MT oxidise (Boden *et al.*, 2011). However, the only one of these enzymes to be positively identified was a DmoAB-type DMS monooxygenase.

Chapters 3 and 4 found that a putative SfnFG-type $DMSO_2$ monooxygenase is encoded by the same *H. sulfonivorans* gene cluster as the DmoAB-type DMS monooxygenase, and that both enzymes are specifically induced in response to $DMSO_2$ as a sole carbon source. As these enzymes are thought to mediate exclusive pathways of MSC metabolism it is puzzling that both enzymes would be encoded so closely together in the genome and expressed at the same time, especially as the disruption of DmoAB prevents the methylotrophic utilisation of $DMSO_2$. Initially, three hypotheses were put forward to resolve this disparity:

1. ***The DMS oxidation pathway is expressed in conjunction with a $DMSO_2$ oxidation pathway.***

2. ***The DMS oxidation pathway is expressed in conjunction with an incomplete $DMSO_2$ oxidization pathway.***

3. ***SfnFG2 functions as an enzyme of the DMS oxidation pathway, such as MT oxidase, DMSO reductase or $DMSO_2$ reductase.***

As described at the end of Chapter 4, none of these are particularly likely given that each hypothesis conflicts with our current understanding of $DMSO_2$ metabolism via the DMS oxidation pathway. However, the functional genomics performed in Chapter 5 on other *Hyphomicrobium* species has now implicated DmoAB and SfnFG in a methylotrophic $DMSO_2$ oxidation pathway, at least in *H. methylovorum* Bras1.

Hydrogen sulfide is predicted to be a major product of the DMS oxidation pathway, supported by the induction of various hydrogen sulfide oxidising enzymes when *H. denitrificans* and *H. methylovorum* were cultivated on DMS as a sole carbon source. In contrast, the cultivation of *H. methylovorum* and *H. sulfonivorans* on $DMSO_2$ as a sole carbon source triggered no such induction versus MeOH, save for a single SoxCD encoded by the *dmoA*/*sfnG* gene cluster, suggesting that $DMSO_2$ metabolism in these species generates sulfite

for oxidation to sulfate via the sulfate adenylylation pathway, rather than the oxidation of thiosulfate via the Sox system.

If DmoAB is actually part of a $DMSO_2$ oxidation pathway instead of a DMS oxidation pathway, this raises a fourth potential hypothesis for $DMSO_2$ utilisation in *H. sulfonivorans*:

4. ***DmoAB functions as an enzyme of a methylotrophic $DMSO_2$ oxidation pathway, containing the $DMSO_2$ monooxygenase SfnFG2.***

If methylotrophic $DMSO_2$ metabolism does not occur via DMSO or DMS intermediates, then this would explain why *H. sulfonivorans* is incapable of utilising DMS or DMSO as a sole carbon source. It would also explain why the functional genomics performed in this study had been unable to identify suitable candidates for either $DMSO_2$ reductase, DMSO reductase or MT oxidase in either the proteomics or transcriptomics data sets, as we would not necessarily expect these enzymes to be expressed or even present in *H. sulfonivorans* if $DMSO_2$ is metabolised via MSA.

This presents the complication of having to reassess the role of DmoAB in MSC metabolism (Boden *et al.*, 2011), as well as the unidentified $DMSO_2$ reductase, DMSO reductase and MT oxidase previously proposed by Borodina *et al.* (2000, 2002). However, this fourth hypothesis would resolve several conflicts between the multi-omics data sets collected during the course of the project and the current model of MSC metabolism in which $DMSO_2$ is oxidised via a hydrogen sulfide producing DMS oxidation pathway.

6.2.1 Revaluating methylotrophic $DMSO_2$ utilisation in *H. sulfonivorans*

If methylotrophic $DMSO_2$ utilisation in *H. sulfonivorans* occurs via a $DMSO_2$ oxidation pathway, then it would replace the DMS oxidation pathway previously proposed for the organism by Borodina *et al.* (2000) and Boden *et al.* (2011). Rather than discounting the previous work performed in the organism, we may now be able to use the additional information gleaned from the genotyping, phenotyping and functional genomics of *Hyphomicrobium* species performed over the course of this project to re-examine the existing experimental data and assess the potential for a methylotrophic pathway of $DMSO_2$ metabolism.

The DMS oxidation pathway was first proposed by Borodina *et al.* (2000) after studying the enzyme activity of *H. sulfonivorans* cell-free extracts when the organism was cultivated on a sole carbon source of $DMSO_2$. By measuring the oxidation of NADH to NAD^+ they were

able to demonstrate the $DMSO_2$, DMSO, DMS and MSA-dependent depletion of NADH, which they ascribed to the activity of NADH-dependent $DMSO_2$ reductase, DMSO reductase, DMS monooxygenase and MSA monooxygenase respectively. An oxygen electrode experiment was also performed which demonstrated that *H. sulfonivorans* lysate mediated MT-dependent oxygen uptake, which Borodina *et al.* (2000) suggested may have been mediated by an NADH-independent MT oxidase. However, the enzyme activity assays for DMS monooxygenase, MSA monooxygenase and MT oxidase were not performed against *H. sulfonivorans* biomass cultivated on a non-MSC control compound (Borodina *et al.*, 2000), so it is unknown to what extent this represents a background level for the organism's lysate.

Furthermore, NADH-oxidation enzyme activity assays only measure the depletion of NADH rather than a particular enzyme activity, so it is possible that the enzyme responsible for DMS, DMSO and $DMSO_2$-dependent NADH oxidation have been misattributed. Indeed, Borodina *et al.* (2000) stated in their original publication that the $DMSO_2$-dependent NADH depletion may be caused by a $DMSO_2$ monooxygenase instead of a $DMSO_2$ reductase and that the MSA-dependent NADH depletion that they observed may be caused by an MSA monooxygenase. If the hypothesis that $DMSO_2$ utilisation occurs via a $DMSO_2$ oxidation is true, then this enzyme activity would be consistent with the proposed activities of SfnFG2 and DmoAB in a methylotrophic $DMSO_2$ oxidation pathway.

Boden *et al.* (2011) later purified the large subunit (DmoA) and small subunit (DmoB) of an $FMNH_2$-dependent monooxygenase that is expressed when *H. sulfonivorans* biomass is cultivated on $DMSO_2$ as a sole carbon source. They went on to demonstrate that the enzyme mediated the DMS-dependent depletion of NADH, improved by the addition of FMN, but found that no NADH-depletion took place for alky thiols, sulfoxides, sulfones or sulfonates. However, the enzyme did mediate NADH-depletion in the presence of some other alkyl sulfides at <50% the specific activity of DMS, such as diethylsulfide and ethylmethylsulfide, suggesting that the enzyme primarily acts on DMS. On this evidence the DmoAB-type monooxygenase was therefore assigned as the DMS monooxygenase of the DMS oxidation pathway that has already been proposed by Borodina *et al.* (2000).

For the methylotrophic $DMSO_2$ oxidation pathway to be a credible alternative to the DMS oxidation pathway, the substrate specificity of DmoAB recorded by Boden *et al.* (2011) will have to be revisited. The DMS monooxygenase activity displayed by DmoAB does not preclude the potential for the enzyme to also act as an MSA monooxygenase, although the enzyme's inability to utilise alkyl sulfonates or alkyl sulfones would suggest that the enzyme does not function as either an MSA or $DMSO_2$ monooxygenase. It is possible that a technical

issue prevented the identification of this activity in the previous set of experiments and that an absence of evidence is not necessarily evidence for absence, but ultimately the only way to confirm DmoAB's activity will be to re-purify and/or express the enzyme to repeat the substrate specificity assays performed by Boden *et al.* (2011).

6.2.2 Conclusions

The methylotrophic pathway of $DMSO_2$ oxidation in *H. sulfonivorans* S1 appears to involve the $FMNH_2$-dependent monooxygenases DmoAB and SfnFG2, encoded by the *dmoA*/*sfnG2* gene cluster (C6Y62_13175:13255), while the alternative pathway is likely to involve the $FMNH_2$-dependent monooxygenases MsuDE and SfnFG1, encoded by the *msuD*/*sfnG1* gene cluster (C6Y62_00835:00920). SfnG1 and SfnG2 have tentatively been assigned the role of $DMSO_2$ monooxygenases, MsuDE as an MSA monooxygenase and it has been proposed that the DMS monooxygenase DmoAB may also function as an MSA monooxygenase.

However, further research is needed to verify the activity of these putative monooxygenases, identify the unknown DMS and DMSO dehydrogenases of the alternative $DMSO_2$ utilisation pathway and establish the identity of the MSA, $DMSO_2$, DMSO and DMS transporters responsible for the uptake of MSCs for assimilatory metabolism.

6.3 Non-methylotrophic $DMSO_2$ oxidation pathway of *H. sulfonivorans* S1

Phenotyping of the DMS monooxygenase gene disruption mutant *ΔdmoA* in Chapter 3 demonstrated that *H. sulfonivorans* S1 contains a non-methylotrophic pathway of MSC metabolism which can mediate the utilisation of $DMSO_2$ as a sulfur source, even in the absence of a functional DMS monooxygenase. A search for MSC degrading enzymes from other organisms identified the large subunits of a putative SfnFG-type $DMSO_2$ monooxygenase (C6Y62_00885) and MsuDE-type MSA monooxygenase (C6Y62_00835) in the same gene cluster, as well as a candidate for an SfnF/MsuE-type FMN reductase (C6Y62_00880).

Functional genomics of $DMSO_2$ metabolism in *H. sulfonivorans* described in Sections 4.2 and 4.3 revealed that all three of these putative SfnF, MsuD and SfnF/MsuE subunits, plus a putative SsuABC-type alkanesulfonate transporter (C6Y62_00870, C6Y62_00865 and C6Y62_00860) encoded by the same gene cluster, are specifically induced in response to the

cultivation of *H. sulfonivorans* on $DMSO_2$ as a sole sulfur source, but repressed when the organism is cultivated on $DMSO_2$ as a sole carbon source in the presence of sulfate.

SfnFG and MsuDE, characterised in *Pseudomonas fluorescens* and *Pseudomonas putida* respectively (Kertesz *et al.*, 1999; Wicht, 2016), have previously been identified as enzymes of a $DMSO_2$ oxidation pathway in *Pseudomonas* species which mediates the utilisation of MSCs as a sulfur source (Endoh *et al.*, 2004). Similarly, SsuABC-type alkanesulfonate transporters have previously been found to mediate the transport of extracellular alkanesulfonates into *Escherichia coli* for use as a sulfur source (Eichhorn *et al.*, 2000). Furthermore, an SsuABC-like arylsulfonate transporter, showing close homology to the aforementioned *E. coli* alkanesulfonate transporter, has also been shown to be encoded in the same gene cluster as the MsuDE of a *Pseudomonas putida* strain, where it has been shown to be associated with the assimilation of arylsulfides as a sulfur source (Endoh *et al.*, 2004).

Returning to the SfnFG1 and MsuDE type $FMNH_2$-dependent monooxygenases of *H. sulfonivorans*, this combination of close homology to known enzymes, gene synteny with similar clusters in other organisms and sulfur-specific induction in response to $DMSO_2$ make it highly likely that these enzymes are involved in the utilisation of MSCs as a sulfur source in S1. It is therefore proposed that SfnG1 and MsuDE are part of an alternative $DMSO_2$ oxidation pathway, induced in response to sulfate starvation, for the utilisation of DMS, DMSO, $DMSO_2$ and MSA as a sole sulfur source in the absence of a functional DmoAB-type monooxygenase, regardless of the role that DmoAB plays in MSC metabolism. However, the proposal of this pathway raises several questions that will need to be answered by future work, discussed below.

Why is the alternative $DMSO_2$ oxidation pathway non-methylotrophic?

Both SfnFG1 and MsuDE of the alternative $DMSO_2$ oxidation pathway are expected to produce formaldehyde, the same C1 intermediate of the methylotrophic MSC degradation pathway, so it is odd that the pathway cannot be used by *H. sulfonivorans* to utilise the compound as a carbon source. One possible explanation is that the pathway is too inefficient to metabolise $DMSO_2$ as a sole carbon source and/or incapable of detoxifying its downstream metabolites, though this may be unlikely if the DmoAB is successfully characterised as an MSA monooxygenase due to both the methylotrophic and non-methylotrophic $DMSO_2$ pathways sharing the same metabolites, formaldehyde and sulfite.

Alternatively, it may be that the pathway does indeed generate formaldehyde from $DMSO_2$ and that this C1 compound can be assimilated as a methylotrophic carbon source, but

that the pathway is not expressed when *H. sulfonivorans* utilises $DMSO_2$ as a sole carbon source in the presence of replete sulfur; meaning the pathway is induced by sulfur starvation and/or repressed by sulfate. It may be possible to test this hypothesis by introducing constitutively active members of the *msuD*/*sfng1* gene cluster into *H. sulfonivorans* Δ*dmoA* to see if this rescues the mutant phenotype; i.e. allows the complemented mutant to utilise $DMSO_2$ as a carbon source.

Why is the alternative $DMSO_2$ oxidation pathway repressed on $DMSO_2$ as carbon/sulfur source?

If this alternative $DMSO_2$ oxidation pathway is involved in the utilisation of $DMSO_2$ as a sulfur source, then we might expect the pathway to be expressed when $DSMO_2$ is utilised by *H. sulfonivorans* as a sole carbon and sulfur source, but this is not the case.

When *H. sulfonivorans* is cultivated on $DMSO_2$ as a sole carbon source it rapidly acidifies the culture media when compared to a non-MSC substrate such as MeOH (Borodina *et al.*, 2000), likely due to the excretion of sulfate into the culture media as seen in other *Hyphomicrobium* species (Koch and Dahl, 2018). As an organism's carbon requirements will be ~30-100 fold higher than its sulfur requirements, the utilisation of $DMSO_2$ as a sole carbon and sulfur source is likely to yield an excess of inorganic sulfur. This means that even in the absence of sulfate as an external sulfur source, sulfate may accumulate within the cell and inhibit the expression of the $DMSO_2$ oxidation pathway, as seen in the proteomics data when *H. sulfonivorans* was cultivated on $DMSO_2$ as a sole carbon and sulfur course.

If this hypothesis is correct, then it may be possible to cultivate the *H. sulfonivorans* Δ*dmoA* mutant on $DMSO_2$ as a sole carbon and sulfur source in continuous culture, removing the excess sulfate produced by methylotrophic $DMSO_2$ metabolism. However, the success of this strategy will be dependent on the level of sulfate required to supresses the *msuD*/*ssuD* gene cluster, and the rate of sulfate efflux from the system exceeding the rate of sulfate production by the *H. sulfonivorans* Δ*dmoA* strain.

What are the unidentified enzymes of DMS and DMSO dehydrogenation?

The utilisation of DMS and DMSO as a sulfur source via a $DMSO_2$ oxidation pathway, would first require their oxidation to $DMSO_2$, potentially mediated by DMS and/or DMSO dehydrogenases or monooxygenases. Although these enzymes have yet to be identified in

Hyphomicrobium species, the *msuD*/*sfnG1* gene cluster may contain several candidates for these enzymes: These are a putative SfnB-type oxidoreductase (C6Y62_00890), cysteine hydrolase (C6Y62_00875), amidohydrolase (C6Y62_00845) and hydroxyacylglutathione hydrolase (C6Y62_00840). If one of these genes is indeed a DMS or DMSO dehydrogenase, we may expect the targeted deletion or disruption of that gene to prevent *H. sulfonivorans* from utilising DMS and/or DMSO as a sole sulfur source, while still allowing it to utilise $DMSO_2$ and MSA as a sulfur source, and $DMSO_2$ as a carbon source.

Alternatively, *H. sulfonivorans* S1 would be expected to express DMS and DMSO dehydrogenases when the strain is cultivated on DMS and/or DMSO as a sole sulfur, at which point it may be possible to detect the expression of these enzymes in the organism's proteome. It may then be possible to identify the enzymes via the fractionation and enzyme assay of *H. sulfonivorans* lysate, searching for either DMS-dependent NADH-oxidation, the NADH-dependent depletion of DMS measured by GC, and/or the NADH-dependent production or depletion of DMSO. Note that if this work is performed in the Δ*dmoA* strain then it would avoid the risk of false positives arising from DmoAB-mediated DMS oxidation. With the identification of a suitable DMS or DMSO degrading fraction, it would then be possible to identify the enzyme's identity by sequencing of the protein and/or performing mass spectrometry.

Which metabolic pathways mediate the utilisation of MSCs as a sulfur source in other Hyphomicrobium *species?*

Of the five organisms examined in this study only *H. sulfonivorans* S1 contains an *msuD*/*sfnFG* gene cluster. Therefore, the other *Hyphomicrobium* species must be using a different pathway to metabolise MSCs as a sole sulfur source from the MsuDE/SfnFG-type $DMSO_2$ oxidation pathway discussed above. It may therefore be interesting to establish whether these species are using their respective methylotrophic MSC pathways to utilise these compounds as a sulfur source, and if they follow the same pattern of substrate utilisation, i.e. metabolism of DMS via a DMS oxidation pathway and metabolism of $DMSO_2$ via a $DMSO_2$ oxidation pathway.

Although proteomics or transcriptomics could be used to answer this question, a simpler solution may be to perform RT-PCR or to measure the expression of the *mtoX*, *sfnG* and *dmoA* genes when these species are cultivated on various MSCs as either a carbon or sulfur source.

6.4 Future research into the MSC metabolism of *Hyphomicrobium* species

The research performed in this project has now opened several new avenues of study in the characterisation of MSC metabolism in *Hyphomicrobium* species. Outlined below is a description of the major gaps in our knowledge and the next steps that could be taken to resolve them.

6.4.1 Characterising the putative $FMNH_2$-dependent MSC monooxygenases

With the discovery that DmoAB-type DMS monooxygenases appear in an MtoX-independent pathway in some *Hyphomicrobium* species, suggesting that the enzyme may play an alternative role in methylotrophic MSC metabolism, the reassessment of this enzyme should be a high priority, as many of the potential putative metabolic pathways proposed in this study depend on the function of this enzyme. Furthermore, functional genomics has implicated several other putative $FMNH_2$-dependent monooxygenases in the metabolism of MSCs by *Hyphomicrobium* species, specifically SfnFG1, SfnFG2, MsuD and LadA from *H. sulfonivorans*, and an SfnFG from *H. methylovorum* Bras1.

The first step would be to purify the predicted monooxygenase large subunits and FMN reductase small subunits of each of these enzymes, as previously performed by Boden *et al.* (2000) in their original characterisation of DmoAB. As we have the nucleotide sequence of each of these enzymes it would either be possible to express these enzymes in *H. sulfonivorans*, extract the proteins from biomass purify via size exclusion chromatography, or to heterologously express them in another model organism such as *E. coli*, potentially adding a polyhistidine-tag for histidine-affinity chromatography.

Based on the phylogenetics of these enzymes and their homology to previously characterised monooxygenases from other species, we may expect them to be NADH-dependent oxidisers of MSCs such as $DMSO_2$, DMSO, DMS, MT and MSA, other aryl or alkyl sulfonates, sulfides and sulfoxides, or in the case of LadA long-chain alkanes. This presents the option of performing chemical assays for aldehyde production, alcohol production, NADH oxidation, hydrogen peroxide production (oxidase rather than monooxygenase activity), oxygen electrode experiments for respiration, or using gas chromatography to measure the depletion or production of volatile sulfur such as compounds DMS, MT and H_2S. Note that many of these techniques have already been applied to the problem of characterising enzymes of MSC metabolism in *Hyphomicrobium* species (Borodina *et al.*, 2000; Borodina *et al.*, 2002; Boden *et al.*, 2011, Eyice *et al.*, 2017).

6.4.2 Targeted mutagenesis of *Hyphomicrobium* species

In addition to characterising the substrate range of the $FMNH_2$-dependent monooxygenases, it would be useful to use targeted mutagenesis to assess the relationship of these enzymes with MSC metabolism in *Hyphomicrobium* species. Although the abundance of potentially redundant $FMNH_2$-dependent monooxygenases in the *Hyphomicrobium* may hinder such an analysis, the characterisation of *ΔdmoA* performed in Chapter 3 demonstrates that disrupting these enzymes has the potential to cause stark differences in MSC metabolism between the different strains; i.e. disrupting *dmoA* prevented methylotrophic growth on $DMSO_2$.

The *ΔdmoA* mutant used in this study, produced by Scanlan and Schäfer, was transformed by replacing an internal sequence of the gene with a gentamycin resistance cassette by homologous recombination. However, Koch and Dahl (2018) have since published a protocol for generating markerless deletion mutants in *H. denitrificans*, which may be a more sustainable platform for producing multiple deletion mutants of *Hyphomicrobium* species; marker insertion mutagenesis requires a different antibiotic for each transformation, limiting the number of possible gene deletions by the availability of viable antibiotics against the bacteria for transformation.

The most pressing targets for mutagenesis in the *Hyphomicrobium* species include the MSC monooxygenase encoding genes *msmA*, *msuD, sfnG1/ 2*, the MT oxidase encoding *mtoX* gene and putative alkane monooxygenase encoding *ladA*. Discovering the impact of each gene's disruption would help to identify their respective biological functions and test the new pathways of MSC metabolism that have been proposed throughout this study. Other targets may include the genes encoding other enzymes of interest from the *msuD*/*sfnG*, *dmoA*/*sfnG* or *mtoX* gene clusters, such as prospective MSC transporters (*ssuABC*, *livFGHKM*, porins), the potential NADH-dependent DMS monooxygenase azoreductase (*azoR*), transcriptional regulators (*cbl*, *lysR*) or candidates for other MSC degrading enzymes (amidohydrolase, cysteine hydrolase, alkane 1-monooxygenase).

6.4.3 Cell-free extract activity assays of *Hyphomicrobium* species

Studying MSC metabolism in the various *Hyphomicrobium* species by protein purification and mutagenesis has the potential to be both time consuming and labour intensive,

making a full characterisation of MSC utilisation in each species unattainable in the short term. However, now armed with a better understanding of the various metabolic processes driving MSC utilisation in *H. denitrificans*, *H. methylovorum* and *H. sulfonivorans*, we may be in a better position to study MSC degradation in these strains by performing enzyme activity assays and oxygen electrode experiments on MSC cultivated cell-free extracts.

For example, there is now strong evidence to suggest that several of the MSC oxidising enzymes, such as MSA monooxygenase and $DMSO_2$ monooxygenase, are $FMNH_2$-dependent enzymes. Work by Boden *et al.* (2011) on the characterisation of DmoA found that the enzyme activity of the enzyme was substantially increased by the addition of an FMN cofactor (~12-fold), so it is possible that the addition of FMN to *Hyphomicrobium* cell lysate may help to stimulate the activity of these putative $FMNH_2$-dependent monooxygenases. Furthermore, functional genomics has demonstrated the specific MSC and carbon/sulfur conditions used to induce the expression of MSC degrading enzymes in *Hyphomicrobium* biomass can have a major influence on which enzymes and processes are induced or repressed: i.e. cultivation of a sole carbon source of DMS leads to MtoX expression, while $DMSO_2$ leads to DmoAB and SnfFG.

Following the previous work performed in *Hyphomicrobium* species, most obvious types of assay to perform would be measuring the production and/or degradation of volatile sulfur compounds by gas chromatography (DMS oxidation pathway) such as DMS, MT and hydrogen sulfide, NADH oxidation activity assays in the presence and absence of cofactors such as FMN or FAD, or assays for the production of other predicted products of MSC metabolism, such as formaldehyde and sulfite.

6.4.4 Studying methylotrophic DMSO metabolism

In Chapter 5, *H. methylovorum* was found to be capable of utilising DMS, DMSO and $DMSO_2$ as a sole carbon source. Functional genomics suggests that *H. methylovorum* utilises DMS a carbon source via a DMS oxidation pathway and $DMSO_2$ via a $DMSO_2$ oxidation pathway, but it is unknown which pathway(s) mediate DMSO metabolism. Discovering this information may also have applications in the characterisation and identification of the unidentified DMS dehydrogenase, DMSO dehydrogenase, DSMO reductase and $DMSO_2$ reductase that may or may not facilitate the oxidation and reduction of dimethyl sulfur compounds.

Although we could use proteomics or transcriptomics to distinguish between the DMS and $DMSO_2$ oxidation pathways, as were previously used in Chapters 3 and 4 to examine other processes of MSC metabolism, it may be simpler and more cost effective to perform a quantitative PCR (RT-PCR) experiment on the extracts of *H. methylovorum* when cultivated on a sole carbon source of either the DMS (DMS oxidation pathway), $DMSO_2$ ($DMSO_2$ oxidation pathway) or DMSO (pathway unknown). Furthermore, it may also be possible to use this same process to examine which of these metabolic pathways is used to assimilate $DMSO_2$, DMSO and DMS assimilation as a sole sulfur source.

6.4.5 Investigating MSA transport in *H. facile* Bras3

The phenotyping of *Hyphomicrobium* species performed in Chapters 3 and 5 showed that both *H. sulfonivorans* and *H. facile* are capable of utilising MSA as a sulfur source and carbon/sulfur source respectively. Although neither species contains an MsmEFGH-type MSA transporter (Jamshad *et al.,* 2006), the capacity of *H. sulfonivorans* to utilise MSA as a sulfur source can be explained by the induction of a putative SsuABC-type alkanesulfonate transporter that appears to be induced in response to sulfate starvation. However, this particular ABC transporter is missing from the *H. denitrificans*, *H. methylovorum* Bras1, *H. facile* Bras3 and *H.* sp. VS genomes, suggesting that another transport system must be facilitating the utilisation of MSA in *H. facile*.

Resource permitting, it may be useful to continue using comparative proteomics of MSC metabolism in *H. facile* or other *Hyphomicrobium* species, cultivated on other substrates such as MSA and DMSO.

6.5 Final conclusions

The overarching goal of the project has been to use functional genomics to improve our understanding of MSC metabolism in the model organism *Hyphomicrobium sulfonivorans* S1, then to apply this information to the study of other *Hyphomicrobium* strains to gain a better understanding of the bacterial processes of assimilatory MSC metabolism. This has led to the significant discovery that *H. sulfonivorans* contains two distinct mechanisms for the utilisation of $DMSO_2$ as either a sole carbon source or sole sulfur source, with strong evidence that both pathways involve the sequential oxidation of $DMSO_2$ to sulfite in a process not previously identified in a *Hyphomicrobium* species.

Furthermore, the information gained in *H. sulfonivorans* has also been successfully applied to the study of MSC metabolism in two more *Hyphomicrobium* species, *H. denitrificans* ATCC51888 and *H. methylovorum* Bras1, identifying an additional pathway of DMS oxidation, raising a new perspective for the activity of DmoAB-type monooxygenases and paving the way for further research into MSC metabolism in the *Hyphomicrobium* genus.

References & Appendixes

References

ADAMS, M. A. & JIA, Z. 2005. Structural and biochemical evidence for an enzymatic quinone redox cycle in *Escherichia coli* identification of a novel quinol monooxygenase. *Journal of Biological Chemistry,* 280**,** 8358-8363.

ALBERS, P., LOOD, C., ÖZTURK, B., HOREMANS, B., LAVIGNE, R., VAN NOORT, V., DE MOT, R., MARCHAL, K., SANCHEZ-RODRIGUEZ, A. & SPRINGAEL, D. 2018a. Catabolic task division between two near-isogenic subpopulations co-existing in a herbicide-degrading bacterial consortium: consequences for the interspecies consortium metabolic model. *Environmental Microbiology,* 20**,** 85-96.

ALBERS, P., WEYTJENS, B., DE MOT, R., MARCHAL, K. & SPRINGAEL, D. 2018b. Molecular processes underlying synergistic linuron mineralization in a triple-species bacterial consortium biofilm revealed by differential transcriptomics. *MicrobiologyOpen,* 7**,** e00559.

ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. & LIPMAN, D. J. 1990. Basic local alignment search tool. *Journal of Molecular Biology,* 215**,** 403-410.

AMARATUNGA, K., GOODWIN, P. M., O'CONNOR, C. D. & ANTHONY, C. 1997. The methanol oxidation genes mxaFJGIR (S) ACKLD in *Methylobacterium extorquens*. *FEMS Microbiology Letters,* 146**,** 31-38.

ANDERS, S. & HUBER, W. 2010. Differential expression analysis for sequence count data. *Genome Biology,* 11**,** R106.

ANGIUOLI, S. V., GUSSMAN, A., KLIMKE, W., COCHRANE, G., FIELD, D., GARRITY, G. M., KODIRA, C. D., KYRPIDES, N., MADUPU, R. & MARKOWITZ, V. 2008. Toward an online repository of Standard Operating Procedures (SOPs) for (meta) genomic annotation. *OMICS A Journal of Integrative Biology,* 12**,** 137-141.

APWEILER, R., BAIROCH, A., WU, C. H., BARKER, W. C., BOECKMANN, B., FERRO, S., GASTEIGER, E., HUANG, H., LOPEZ, R. & MAGRANE, M. 2004. UniProt: the universal protein knowledgebase. *Nucleic Acids Research,* 32**,** D115-D119.

ATTWOOD, M. M. & HARDER, W. 1972. A rapid and specific enrichment procedure for *Hyphomicrobium* spp. *Antonie van Leeuwenhoek,* 38**,** 369-377.

ATTWOOD, M. M. & HARDER, W. 1978. Formate assimilation by *Hyphomicrobium* X. *FEMS Microbiology Letters,* 3**,** 111-114.

AWANO, S., KOSHIMUNE, S., KURIHARA, E., GOHARA, K., SAKAI, A., SOH, I., HAMASAKI, T., ANSAI, T. & TAKEHARA, T. 2004. The assessment of methyl mercaptan, an important clinical marker for the diagnosis of oral malodor. *Journal of Dentistry,* 32**,** 555-559.

AYERS, G. P. & GILLETT, R. W. 2000. DMS and its oxidation products in the remote marine atmosphere: implications for climate and atmospheric chemistry. *Journal of Sea Research,* 43**,** 275-286.

BANKEVICH, A., NURK, S., ANTIPOV, D., GUREVICH, A. A., DVORKIN, M., KULIKOV, A. S., LESIN, V. M., NIKOLENKO, S. I., PHAM, S. & PRJIBELSKI, A. D. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology,* 19**,** 455-477.

BAXTER, N. J., SCANLAN, J., DE MARCO, P., WOOD, A. P. & MURRELL, J. C. 2002. Duplicate copies of genes encoding methanesulfonate monooxygenase in *Marinosulfonomonas methylotropha* strain TR3 and detection of methanesulfonate utilizers in the environment. *Applied and Environmental Microbiology,* 68**,** 289-296.

BENTLEY, R. & CHASTEEN, T. G. 2004. Environmental VOSCs—formation and degradation of dimethyl sulfide, methanethiol and related materials. *Chemosphere,* 55**,** 291-317.

BODEN, R., BORODINA, E., WOOD, A. P., KELLY, D. P., MURRELL, J. C. & SCHÄFER, H. 2011a. Purification and characterization of dimethylsulfide monooxygenase from *Hyphomicrobium sulfonivorans*. *Journal of Bacteriology*, 193, 1250-8.

BODEN, R., KELLY, D. P., MURRELL, J. C. & SCHÄFER, H. 2010. Oxidation of dimethylsulfide to tetrathionate by Methylophaga thiooxidans sp. nov.: a new link in the sulfur cycle. *Environmental microbiology,* 12**,** 2688-2699.

BODEN, R., MURRELL, J. C. & SCHÄFER, H. 2011b. Dimethylsulfide is an energy source for the heterotrophic marine bacterium *Sagittula stellata*. *FEMS Microbiology Letters,* 322**,** 188-193.

BOLGER, A. M., LOHSE, M. & USADEL, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics,* 30**,** 2114-2120.

BORATYN, G. M., CAMACHO, C., COOPER, P. S., COULOURIS, G., FONG, A., MA, N., MADDEN, T. L., MATTEN, W. T., MCGINNIS, S. D. & MEREZHUK, Y. 2013. BLAST: a more efficient report with usability improvements. *Nucleic Acids Research,* 41**,** W29-W33.

BORODINA, E., KELLY, D. P., RAINEY, F. A., WARD-RAINEY, N. L. & WOOD, A. P. 2000. Dimethylsulfone as a growth substrate for novel methylotrophic species of *Hyphomicrobium* and *Arthrobacter*. *Archives of Microbiology,* 173**,** 425-437.

BORODINA, E., KELLY, D. P., SCHUMANN, P., RAINEY, F. A., WARD-RAINEY, N. L. & WOOD, A. P. 2002. Enzymes of dimethylsulfone metabolism and the phylogenetic characterization of the facultative methylotrophs *Arthrobacter sulfonivorans* sp. nov., *Arthrobacter methylotrophus* sp. nov., and *Hyphomicrobium sulfonivorans* sp. nov. *Archives of Microbiology,* 177**,** 173-183.

BOURGON, R., GENTLEMAN, R. & HUBER, W. 2010. Independent filtering increases detection power for high-throughput experiments. *Proceedings of the National Academy of Sciences,* 107**,** 9546-9551.

BROOKE, A., DUCHARS, M. & ATTWOOD, M. M. 1987. Nitrogen assimilation in the facultative methylotroph *Hyphomicrobium* X. *FEMS Microbiology Letters,* 41**,** 41-45.

BROWN, P. J., KYSELA, D. T., BUECHLEIN, A., HEMMERICH, C. & BRUN, Y. V. 2011. Genome sequences of eight morphologically diverse Alphaproteobacteria. *Journal of Bacteriology***,** JB. 05453-11.

CAMACHO, C., COULOURIS, G., AVAGYAN, V., MA, N., PAPADOPOULOS, J., BEALER, K. & MADDEN, T. L. 2009. BLAST+: architecture and applications. *BMC Bioinformatics,* 10**,** 421.

CANNON, R. J. & HO, C.-T. 2018. Volatile sulfur compounds in tropical fruits. *Journal of Food and Drug Analysis*.

CARRIÓN, O., CURSON, A., KUMARESAN, D., FU, Y., LANG, A., MERCADÉ, E. & TODD, J. 2015. A novel pathway producing dimethylsulphide in bacteria is widespread in soil environments. *Nature Communications,* 6**,** 6579.

CARRIÓN, O., PRATSCHER, J., CURSON, A. R., WILLIAMS, B. T., ROSTANT, W. G., MURRELL, J. C. & TODD, J. D. 2017. Methanethiol-dependent dimethylsulfide production in soil environments. *The ISME Journal,* 11**,** 2379.

CHARLSON, R. J., LOVELOCK, J. E., ANDREAE, M. O. & WARREN, S. G. 1987. Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. *Nature,* 326**,** 655.

CHEN, N. H., DJOKO, K. Y., VEYRIER, F. J. & MCEWAN, A. G. 2016. Formaldehyde stress responses in bacterial pathogens. *Frontiers in Microbiology,* 7**,** 257.

CHEVENET, F., BRUN, C., BAÑULS, A.-L., JACQ, B. & CHRISTEN, R. 2006. TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics,* 7**,** 439.

CHISTOSERDOVA, L. 2011. Modularity of methylotrophy, revisited. *Environmental Microbiology,* 13**,** 2603-2622.

CHISTOSERDOVA, L., CHEN, S.-W., LAPIDUS, A. & LIDSTROM, M. E. 2003. Methylotrophy in *Methylobacterium extorquens* AM1 from a genomic point of view. *Journal of Bacteriology,* 185**,** 2980-2987.

CHISTOSERDOVA, L., KALYUZHNAYA, M. G. & LIDSTROM, M. E. 2009. The expanding world of methylotrophic metabolism. *Annual Review of Microbiology,* 63**,** 477-499.

CHONGCHAROEN, R., SMITH, T. J., FLINT, K. P. & DALTON, H. 2005. Adaptation and acclimatization to formaldehyde in methylotrophs capable of high-concentration formaldehyde detoxification. *Microbiology,* 151**,** 2615-2622.

CHU, F. & LIDSTROM, M. E. 2016. XoxF acts as the predominant methanol dehydrogenase in the type I methanotroph *Methylomicrobium buryatense*. *Journal of bacteriology***,** JB. 00959-15.

CLAUS, M. T., ZOCHER, G. E., MAIER, T. H. & SCHULZ, G. E. 2005. Structure of the O-acetylserine sulfhydrylase isoenzyme CysM from *Escherichia coli*. *Biochemistry,* 44**,** 8620-6.

COLBY, J. & ZATMAN, L. 1973. Trimethylamine metabolism in obligate and facultative methylotrophs. *Biochemical Journal,* 132**,** 101-112.

CONSORTIUM, U. 2018. UniProt: the universal protein knowledgebase. *Nucleic acids Research,* 46**,** 2699.

CROWTHER, G. J., KOSÁLY, G. & LIDSTROM, M. E. 2008. Formate as the main branch point for methylotrophic metabolism in *Methylobacterium extorquens* AM1. *Journal of Bacteriology,* 190**,** 5057-5062.

CURSON, A., LIU, J., BERMEJO MARTÍNEZ, A., GREEN, R., CHAN, Y., CARRIÓN, O., WILLIAMS, B., ZHANG, S., YANG, G., BULMAN PAGE, P., ZHANG, X. & TODD, J. 2017. Dimethylsulfoniopropionate biosynthesis in marine bacteria and identification of the key gene in this process. *Nature Microbiology,* 2**,** 17009.

CURSON, A., ROGERS, R., TODD, J., BREARLEY, C. & JOHNSTON, A. 2008. Molecular genetic analysis of a dimethylsulfoniopropionate lyase that liberates the climate-changing gas dimethylsulfide in several marine α-proteobacteria and *Rhodobacter sphaeroides*. *Environmental Microbiology,* 10**,** 757-767.

CURSON, A. R., TODD, J. D., SULLIVAN, M. J. & JOHNSTON, A. W. 2011. Catabolism of dimethylsulphoniopropionate: microorganisms, enzymes and genes. *Nature Reviews Microbiology,* 9**,** 849.

DACEY, J. W. H., KING, G. M. & WAKEHAM, S. G. 1987. Factors controlling emission of dimethylsulphide from salt marshes. *Nature* 330 643-645.

DE BONT, J., VAN DIJKEN, J. & HARDER, W. 1981. Dimethyl sulphoxide and dimethyl sulphide as a carbon, sulphur and energy source for growth of *Hyphomicrobium* S. *Microbiology,* 127**,** 315-323.

DE MARCO, P., MORADAS-FERREIRA, P., HIGGINS, T. P., MCDONALD, I., KENNA, E. M. & MURRELL, J. C. 1999. Molecular analysis of a novel methanesulfonic acid monooxygenase from the methylotroph *Methylosulfonomonas methylovora*. *Journal of Bacteriology,* 181**,** 2244-2251.

DE ZWART, J. M., NELISSE, P. N. & KUENEN, J. G. 1996. Isolation and characterization of *Methylophaga sulfidovorans* sp. nov.: an obligately methylotrophic, aerobic, dimethylsulfide oxidizing bacterium from a microbial mat. *FEMS Microbiology Ecology,* 20**,** 261-270.

DEREEPER, A., GUIGNON, V., BLANC, G., AUDIC, S., BUFFET, S., CHEVENET, F., DUFAYARD, J.-F., GUINDON, S., LEFORT, V. & LESCOT, M. 2008. Phylogeny. fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research,* 36**,** W465-W469.

DORONINA, N., SOKOLOV, A. & TROTSENKO, Y. A. 1996. Isolation and initial characterization of aerobic chloromethane-utilizing bacteria. *FEMS Microbiology Letters,* 142**,** 179-183.

EDGAR, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research,* 32**,** 1792-1797.

EICHHORN, E. & LEISINGER, T. 2001. *Escherichia coli* utilizes methanesulfonate and L-cysteate as sole sulfur sources for growth. *FEMS Microbiology Letters,* 205**,** 271-275.

EICHHORN, E., VAN DER PLOEG, J. R. & LEISINGER, T. 1999. Characterization of a two-component alkanesulfonate monooxygenase from *Escherichia coli*. *Journal of Biological Chemistry,* 274**,** 26639-26646.

EICHHORN, E., VAN DER PLOEG, J. R. & LEISINGER, T. 2000. Deletion analysis of the Escherichia coli taurine and alkanesulfonate transport systems. *Journal of Bacteriology,* 182**,** 2687-2695.

ENDOH, T., KASUGA, K., HORINOUCHI, M., YOSHIDA, T., HABE, H., NOJIRI, H. & OMORI, T. 2003. Characterization and identification of genes essential for dimethyl sulfide utilization in *Pseudomonas putida* strain DS1. *Applied Microbiology and Biotechnology,* 62**,** 83-91.

ERICKSON III, D., GHAN, S. & PENNER, J. 1990. Global ocean-to-atmosphere dimethyl sulfide flux. *Journal of Geophysical Research*, 95, 7543–7552.

ESCALANTE-SEMERENA, J., RINEHART, K. & WOLFE, R. 1984. Tetrahydromethanopterin, a carbon carrier in methanogenesis. *Journal of Biological Chemistry,* 259**,** 9447-9455.

EVANS, C., KADNER, S. V., DARROCH, L. J., WILSON, W. H., LISS, P. S. & MALIN, G. 2007. The relative significance of viral lysis and microzooplankton grazing as pathways of dimethylsulfoniopropionate (DMSP) cleavage: an *Emiliania huxleyi* culture study. *Limnology and Oceanography,* 52**,** 1036-1045.

EYICE, Ö., MYRONOVA, N., POL, A., CARRIÓN, O., TODD, J. D., SMITH, T. J., GURMAN, S. J., CUTHBERTSON, A., MAZARD, S., MENNINK-KERSTEN, M. A., BUGG, T. D., KRISTOFFER ANDERSSON, K., JOHNSTON, W. B., OP DEN

CAMP & H. J., SCHÄFER, H., 2017. Bacterial SBP56 identified as a Cu-dependent methanethiol oxidase widely distributed in the biosphere. *The ISME Journal,* 12**,** 145.

EYICE, Ö. & SCHÄFER, H. 2016. Culture-dependent and culture-independent methods reveal diverse methylotrophic communities in terrestrial environments. *Archives of Microbiology,* 198**,** 17-26.

FENG, L., WANG, W., CHENG, J., REN, Y., ZHAO, G., GAO, C., TANG, Y., LIU, X., HAN, W. & PENG, X. 2007. Genome and proteome of long-chain alkane degrading *Geobacillus thermodenitrificans* NG80-2 isolated from a deep-subsurface oil reservoir. *Proceedings of the National Academy of Sciences,* 104**,** 5602-5607.

FERNÁNDEZ, H. T., RICO, I. R., DE LA PRIDA, J. J. & VAN LANGENHOVE, H. 2013. Dimethyl sulfide biofiltration using immobilized *Hyphomicrobium* VS and *Thiobacillus thioparus* TK-m in sugarcane bagasse. *Environmental Technology,* 34**,** 257-262.

FRIEDRICH, C. G., BARDISCHEWSKY, F., ROTHER, D., QUENTMEIER, A. & FISCHER, J. 2005b. Prokaryotic sulfur oxidation. *Current Opinion in Microbiology,* 8**,** 253-259.

FUSE, H., TAKIMURA, O., MURAKAMI, K., YAMAOKA, Y. & OMORI, T. 2000. Utilization of dimethyl sulfide as a sulfur source with the aid of light by *Marinobacterium* sp. strain DMS-S1. *Applied and Environmental Microbiology,* 66**,** 5527-5532.

GARCÍA-RAMÍREZ, J. J., SANTOS, M. A. & REVUELTA, J. L. 1995. The *Saccharomyces cerevisiae* RIB4 gene codes for 6, 7-dimethyl-8-ribityllumazine synthase involved in riboflavin biosynthesis molecular characterization of the gene and purification of the encoded protein. *Journal of Biological Chemistry,* 270**,** 23801-23807.

GENG, C. & MU, Y. 2004. Carbonyl sulfide and dimethyl sulfide exchange between lawn and the atmosphere. *Journal of Geophysical Research: Atmospheres,* 109.

GHOSH, W. & DAM, B. 2009. Biochemistry and molecular biology of lithotrophic sulfur oxidation by taxonomically and ecologically diverse bacteria and archaea. *FEMS Microbiology Reviews*, 33, 999–1043.

GIRI, B. S., GOSWAMI, M., PANDEY, R. & KIM, K. 2015. Kinetics and biofiltration of dimethyl sulfide emitted from P&P industry. *Biochemical Engineering Journal,* 102**,** 108-114.

GÖRISCH, H. 2003. The ethanol oxidation system and its regulation in *Pseudomonas aeruginosa*. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics,* 1647**,** 98-102.

GRABARCZYK, D. B. & BERKS, B. C. 2017. Intermediates in the Sox sulfur oxidation pathway are bound to a sulfane conjugate of the carrier protein SoxYZ. *PloS One,* 12**,** e0173395.

GRABARCZYK, D. B., CHAPPELL, P. E., EISEL, B., JOHNSON, S., LEA, S. M. & BERKS, B. C. 2015. Mechanism of thiosulfate oxidation in the SoxA family of cysteine-ligated cytochromes. *Journal of Biological Chemistry***,** M114. 618025.

GUINDON, S., DUFAYARD, J.-F., LEFORT, V., ANISIMOVA, M., HORDIJK, W. & GASCUEL, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology,* 59**,** 307-321.

GUNNISON, A. 1981. Sulphite toxicity: a critical review of in vitro and in vivo data. *Food and Cosmetics Toxicology,* 19**,** 667-682.

GUREVICH, A., SAVELIEV, V., VYAHHI, N. & TESLER, G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics,* 29**,** 1072-1075.

HAFT, D. H., DICUCCIO, M., BADRETDIN, A., BROVER, V., CHETVERNIN, V., O'NEILL, K., LI, W., CHITSAZ, F., DERBYSHIRE, M. K. & GONZALES, N. R. 2017. RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Research,* 46**,** D851-D860.

HARDER, W. & ATTWOOD, M. M. 1975. Oxidation of organic C1 compounds by *Hyphomicrobium* spp. *Antonie van Leeuwenhoek,* 41**,** 421-429.

HARDER, W., ATTWOOD, M. M. & QUAYLE, J. 1973. Methanol assimilation by *Hyphomicrobium* sp. *Microbiology,* 78**,** 155-163.

HENRIQUES, A. C. & DE MARCO, P. 2015. Methanesulfonate (MSA) catabolic genes from marine and estuarine bacteria. *PloS One,* 10**,** e0125735.

HERZ, S., EBERHARDT, S. & BACHER, A. 2000. Biosynthesis of riboflavin in plants. The ribA gene of *Arabidopsis thaliana* specifies a bifunctional GTP cyclohydrolase II/3, 4-dihydroxy-2-butanone 4-phosphate synthase. *Phytochemistry,* 53**,** 723-731.

HEUNISCH, G. W. 1977. Stoichiometry of the reaction of sulfites with hydrogen sulfide ion. *Inorganic Chemistry,* 16**,** 1411-1413.

HIRSCH, P. & CONTI, S. 1964. Biology of budding bacteria. *Archiv für Mikrobiologie,* 48**,** 339-357.

HOFFMANN, E. H., TILGNER, A., SCHRÖDNER, R., BRÄUER, P., WOLKE, R. & HERRMANN, H. 2016. An advanced modeling study on the impacts and atmospheric implications of multiphase dimethyl sulfide chemistry. *Proceedings of the National Academy of Sciences,* 113**,** 11776-11781.

HOPKINS, F. E., BELL, T. G., YANG, M., SUGGETT, D. J. & STEINKE, M. 2016. Air exposure of coral is a significant source of dimethylsulfide (DMS) to the atmosphere. *Scientific Reports,* 6**,** 36031.

HORINOUCHI, M., KASUGA, K., NOJIRI, H., YAMANE, H. & OMORI, T. 1997. Cloning and characterization of genes encoding an enzyme which oxidizes dimethyl sulfide in *Acinetobacter* sp. strain 20B. *FEMS Microbiology Letters,* 155**,** 99-105.

HORINOUCHI, M., YOSHIDA, T., NOJIRI, H., YAMANE, H. & OMORI, T. 1999. Polypeptide requirement of multicomponent monooxygenase DsoABCDEF for dimethyl sulfide oxidizing activity. *Bioscience, Biotechnology, and Biochemistry,* 63**,** 1765-1771.

HOSHINO, T. & KOSE, K. 1990. Cloning, nucleotide sequences, and identification of products of the *Pseudomonas aeruginosa* PAO bra genes, which encode the high-affinity branched-chain amino acid transport system. *Journal of Bacteriology,* 172**,** 5531-5539.

HOWARD, E. C., SUN, S., BIERS, E. J. & MORAN, M. A. 2008. Abundant and diverse bacteria involved in DMSP degradation in marine surface waters. *Environmental Microbiology,* 10**,** 2397-2410.

INOUE, H., INAGAKI, K., ERIGUCHI, S.-I., TAMURA, T., ESAKI, N., SODA, K. & TANAKA, H. 1997. Molecular characterization of the mde operon involved in L-

methionine catabolism of *Pseudomonas putida*. *Journal of Bacteriology,* 179**,** 3956-3962.

ITO, K., TAKAHASHI, M., YOSHIMOTO, T. & TSURU, D. 1994. Cloning and high-level expression of the glutathione-independent formaldehyde dehydrogenase gene from *Pseudomonas putida*. *Journal of Bacteriology,* 176**,** 2483-2491.

JAMSHAD, M., DE MARCO, P., PACHECO, C., HANCZAR, T. & MURREL, J. 2006. Identification, mutagenesis, and transcriptional analysis of the methanesulfonate transport operon of *Methylosulfonomonas methylovora*. *Applied and Environmental Microbiology*, 72, 276-283.

JAP, B. K. & WALIAN, P. J. 1996. Structure and functional mechanism of porins. *Physiological Reviews,* 76**,** 1073-1088.

JARDINE, K., YAÑEZ-SERRANO, A. M., WILLIAMS, J., KUNERT, N., JARDINE, A., TAYLOR, T., ABRELL, L., ARTAXO, P., GUENTHER, A., HEWITT, C. N., HOUSE, E., FLORENTINO, A. P., MANZI, A., HIGUCHI, N., KESSELMEIER, J., BEHRENDT, T., VERES, P. R., DERSTROFF, B., FUENTES, J. D., MARTIN, S. T. & ANDREAE, M. O. 2015. Dimethyl sulfide in the Amazon rain forest. *Global Biogeochemical Cycles,* 29**,** 19-32.

JONKERS, H. M., VAN DER MAAREL, M. J., VAN GEMERDEN, H. & HANSEN, T. A. 1996. Dimethylsulfoxide reduction by marine sulfate-reducing bacteria. *FEMS Microbiology Letters,* 136**,** 283-287.

KAHNERT, A., VERMEIJ, P., WIETEK, C., JAMES, P., LEISINGER, T. & KERTESZ, M. A. 2000. The ssu Locus Plays a Key Role in Organosulfur Metabolism in *Pseudomonas putida* S-313. *Journal of bacteriology,* 182**,** 2869-2878.

KANEHISA, M., SATO, Y., KAWASHIMA, M., FURUMICHI, M. & TANABE, M. 2015. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research,* 44**,** D457-D462.

KANEHISA, M., SATO, Y. & MORISHIMA, K. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of molecular biology,* 428**,** 726-731.

KAPPLER, U. 2011. Bacterial sulfite-oxidizing enzymes. *Biochimica et Biophysica Acta (BBA)-Bioenergetics,* 1807**,** 1-10.

KAPPLER, U. & DAHL, C. 2001. Enzymology and molecular biology of prokaryotic sulfite oxidation. *FEMS Microbiology Letters,* 203**,** 1-9.

KELLER, M., BELLOWS, W. and GUILLARD, R. 1989. Dimethyl sulfide production in marine phytoplankton. *Biogenic Sulfur in the Environment*. 167-182.

KELLY, D. P. & MURRELL, J. C. 1999. Microbial metabolism of methanesulfonic acid. *Archives of Microbiology,* 172**,** 341-348.

KELTJENS, J., POL, A., REIMANN, J. & H. OP DEN CAMP. 2014. PQQ-dependent methanol dehydrogenases: rare-earth elements make a difference. *Applied Microbiology and Biotechnology*, 96, 6163–6183.

KERTESZ, M. A. 2000. Riding the sulfur cycle–metabolism of sulfonates and sulfate esters in Gram-negative bacteria. *FEMS Microbiology Reviews,* 24**,** 135-175.

KERTESZ, M. A., SCHMIDT-LARBIG, K. & WÜEST, T. 1999. A novel reduced flavin mononucleotide-dependent methanesulfonate sulfonatase encoded by the sulfur-

regulated *msu* operon of *Pseudomonas aeruginosa*. *Journal of Bacteriology,* 181**,** 1464-1473.

KETTLE, A. J. & ANDREAE, M. O. 2000. Flux of dimethylsulfide from the oceans: A comparison of updated data sets and flux models. *Journal of Geophysical Research Atmospheres,* 105**,** 793-26.

KIENE, R. P. & LINN, L. J. 2000. The fate of dissolved dimethylsulfoniopropionate (DMSP) in seawater: tracer studies using ^{35}S-DMSP. *Geochimica et Cosmochimica Acta*. 2797-2810.

KIENE, R. P., LINN, L. J. & BRUTON, J. A. 2000. New and important roles for DMSP in marine microbial communities. *Journal of Sea Research,* 43**,** 209-224.

KIENE, R. P. 1988. Dimethyl sulfide metabolism in salt marsh sediments. *FEMS Microbiology Letters,* 53**,** 71-78.

KIENE, R. P. 1996. Production of methanethiol from dimethylsulfoniopropionate in marine surface waters. *Marine Chemistry,* 54**,** 69-83.

KIENE, R. P. & VISSCHER, P. T. 1987. Production and fate of methylated sulfur compounds from methionine and dimethylsulfoniopropionate in anoxic salt marsh sediments. *Applied and Environmental Microbiology*, 53**,** 2426-2434.

KINO, K., MURAKAMI-NITTA, T., OISHI, M., ISHIGURO, S. & KIRIMURA, K. 2004. Isolation of dimethyl sulfone-degrading microorganisms and application to odorless degradation of dimethyl sulfoxide. *Journal of Bioscience and Bioengineering*, 97**,** 82-84.

KOCH, T. & DAHL, C. 2018. A novel bacterial sulfur oxidation pathway provides a new link between the cycles of organic and inorganic sulfur compounds. *The ISME Journal*, 12**,** 2479-2491.

KORHONEN, H., CARSLAW, K., SPRACKLEN, D., MANN, G. & WOODHOUSE, M. 2008. Influence of oceanic dimethyl sulfide emissions on cloud condensation nuclei concentrations and seasonality over the remote Southern Hemisphere oceans: A global model study. *Journal of Geophysical Research*, 113, D15204.

KREDICH, N. M. 2008. Biosynthesis of cysteine. *EcoSal Plus,* 3.

LANGMEAD, B. & SALZBERG, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods,* 9**,** 357.

LEUSTEK, T. 2002. Sulfate metabolism. *The Arabidopsis Book/American Society of Plant Biologists,* 1.

LEVASSEUR, M., GOSSELIN, M. & MICHAUD, S. 1994. A new source of dimethylsulfide (DMS) for the arctic atmosphere: ice diatoms. *Marine Biology,* 121**,** 381-387.

LI, H., HANDSAKER, B., WYSOKER, A., FENNELL, T., RUAN, J., HOMER, N., MARTH, G., ABECASIS, G. & DURBIN, R. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics,* 25**,** 2078-2079.

LI, L., LIU, X., YANG, W., XU, F., WANG, W., FENG, L., BARTLAM, M., WANG, L. & RAO, Z. 2008. Crystal structure of long-chain alkane monooxygenase (LadA) in complex with coenzyme FMN: unveiling the long-chain alkane hydroxylase. *Journal of Molecular Biology,* 376**,** 453-465.

LIDBURY, I., KRÖBER, E., ZHANG, Z., ZHU, Y., MURRELL, J. C., CHEN, Y. & SCHÄFER, H. 2016. A mechanism for bacterial transformation of dimethylsulfide to

dimethylsulfoxide: a missing link in the marine organic sulfur cycle. *Environmental Microbiology,* 18**,** 2754-2766.

LIEBEKE, M., PÖTHER, D. C., VAN DUY, N., ALBRECHT, D., BECHER, D., HOCHGRÄFE, F., LALK, M., HECKER, M. & ANTELMANN, H. 2008. Depletion of thiol-containing proteins in response to quinones in *Bacillus subtilis*. *Molecular Microbiology,* 69**,** 1513-1529.

LIU, G., ZHOU, J., FU, Q. S. & WANG, J. 2009. The *Escherichia coli* azoreductase AzoR is involved in resistance to thiol-specific stress caused by electrophilic quinones. *Journal of Bacteriology,* 191**,** 6394-6400.

LIU, H., SHI, C., WU, T., JIA, Q., ZHAO, J. & WANG, X. 2017. Isolation and characterization of methanethiol-producing bacteria from agricultural soils. *Pedosphere,* 27**,** 1083-1091.

LOMANS, B. P., MAAS, R., LUDERER, R., DEN CAMP, H. J. O., POL, A., VAN DER DRIFT, C. & VOGELS, G. D. 1999. Isolation and characterization of *Methanomethylovorans hollandica* gen. nov., sp. nov., isolated from freshwater sediment, a methylotrophic methanogen able to grow on dimethyl sulfide and methanethiol. *Applied and Environmental Microbiology,* 65**,** 3641-3650.

LONG, Y., ZHANG, S., FANG, Y., DU, Y., LIU, W., FANG, C. & SHEN, D. 2017. Dimethyl sulfide emission behavior from landfill site with air and water control. *Biodegradation,* 28**,** 327-335.

LÜ, C., XIA, Y., LIU, D., ZHAO, R., GAO, R., LIU, H. & XUN, L. 2017. *Cupriavidus necator* H16 uses flavocytochrome c-sulfide dehydrogenase to oxidize self-produced and spiked sulfide. *Applied and Environmental Microbiology***,** 83, 1610-17.

MARX, C. J., CHISTOSERDOVA, L. & LIDSTROM, M. E. 2003. Formaldehyde-detoxifying role of the tetrahydromethanopterin-linked pathway in *Methylobacterium extorquens* AM1. *Journal of Bacteriology,* 185**,** 7160-7168.

MATRAI, P. & KELLER, M. 1993. Dimethylsulfide in a large-scale coccolithophore bloom in the Gulf of Maine. *Continental Shelf Research,* 13**,** 831-843.

MATRAI, P. & KELLER, M. 1994. Total organic sulfur and dimethylsulfoniopropionate in marine phytoplankton: intracellular variations. *Marine Biology,* 119**,** 61-68.

MCADAMS, H. H. 2006. Bacterial stalks are nutrient-scavenging antennas. *Proceedings of the National Academy of Sciences,* 103**,** 11435-11436.

MCANULLA, C., WOODALL, C. A., MCDONALD, I. R., STUDER, A., VUILLEUMIER, S., LEISINGER, T. & MURRELL, J. C. 2001. Chloromethane utilization gene cluster from *Hyphomicrobium chloromethanicum* strain CM2T and development of functional gene probes to detect halomethane-degrading bacteria. *Applied and Environmental Microbiology,* 67**,** 307-316.

MCDEVITT, C. A., HUGENHOLTZ, P., HANSON, G. R. & MCEWAN, A. G. 2002. Molecular analysis of dimethyl sulphide dehydrogenase from *Rhodovulum sulfidophilum*: its place in the dimethyl sulphoxide reductase family of microbial molybdopterin-containing enzymes. *Molecular Microbiology,* 44**,** 1575-1587.

MCDONALD, I. R., DORONINA, N. V., TROTSENKO, Y. A., MCANULLA, C. & MURRELL, J. C. 2001. *Hyphomicrobium chloromethanicum* sp. nov. and *Methylobacterium chloromethanicum* sp. nov., chloromethane-utilizing bacteria isolated from a polluted environment. *International Journal of Systematic and Evolutionary Microbiology,* 51**,** 119-122.

MEIBERG, J. & HARDER, W. 1978. Aerobic and anaerobic metabolism of trimethylamine, dimethylamine and methylamine in *Hyphomicrobium* X. *Microbiology,* 106**,** 265-276.

MILLER, B. L., WILLIAMS, T. D. & SCHÖNEICH, C. 1996. Mechanism of sulfoxide formation through reaction of sulfur radical cation complexes with superoxide or hydroxide ion in oxygenated aqueous solution. *Journal of the American Chemical Society,* 118**,** 11014-11025.

MOORE, R. L. 1981. The biology of *Hyphomicrobium* and other prosthecate, budding bacteria. *Annual Reviews in Microbiology,* 35**,** 567-594.

MYRONOVA, N., KITMITTO, A., COLLINS, R. F., MIYAJI, A. & DALTON, H. 2006. Three-dimensional structure determination of a protein supercomplex that oxidizes methane to formaldehyde in *Methylococcus capsulatus* (Bath). *Biochemistry,* 45**,** 11905-11914.

NAKAGAWA, T., MITSUI, R., TANI, A., SASA, K., TASHIRO, S., IWAMA, T., HAYAKAWA, T. & KAWAI, K. 2012. A catalytic role of XoxF1 as La3+-dependent methanol dehydrogenase in *Methylobacterium extorquens* strain AM1. *PloS One,* 7**,** e50480.

NAKANISHI, M., YATOME, C., ISHIDA, N. & KITADE, Y. 2001. Putative ACP phosphodiesterase gene (acpD) encodes an azoreductase. *Journal of Biological Chemistry*.

NAKATANI, T., OHTSU, I., NONAKA, G., WIRIYATHANAWUDHIWONG, N., MORIGASAKI, S. & TAKAGI, H. 2012. Enhancement of thioredoxin/glutaredoxin-mediated L-cysteine synthesis from S-sulfocysteine increases L-cysteine production in *Escherichia coli*. *Microbial Cell Factories,* 11**,** 62.

NEVITT, G. A. 2000. Olfactory foraging by Antarctic procellariiform seabirds: life at high Reynolds numbers. *The Biological Bulletin,* 198**,** 245-253.

NI, S. & BOONE, D. R. 1991. Isolation and characterization of a dimethyl sulfide-degrading methanogen, *Methanolobus siciliae* HI350, from an oil well, characterization of *M. siciliae* T4/MT, and emendation of *M. siciliae*. *International Journal of Systematic and Evolutionary Microbiology,* 41**,** 410-416.

OMORI, T., SAIKI, Y., KASUGA, K. & KODAMA, T. 1995. Desulfurization of alkyl and aromatic sulfides and sulfonates by dibenzothiophene-desulfurizing *Rhodococcus* sp. strain SY1. *Bioscience, Biotechnology, and Biochemistry,* 59**,** 1195-1198.

PARKER, J. K. 2015. Introduction to aroma compounds in foods. *Flavour Development, Analysis and Perception in Food and Beverages, Elsevier,* 3-30.

PERRAUD, V., HORNE, J. R., MARTINEZ, A. S., KALINOWSKI, J., MEINARDI, S., DAWSON, M. L., WINGEN, L. M., DABDUB, D., BLAKE, D. R. & GERBER, R. B. 2015. The future of airborne sulfur-containing particles in the absence of fossil fuel sulfur dioxide emissions. *Proceedings of the National Academy of Sciences,* 112**,** 13514-13519.

PERTEA, M., KIM, D., PERTEA, G. M., LEEK, J. T. & SALZBERG, S. L. 2016. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols,* 11**,** 1650.

PINTO, R., TANG, Q., BRITTON, W., LEYH, T. & TRICCAS, J. 2004. The *Mycobacterium tuberculosis cysD* and *cysNC* genes form a stress-induced operon that encodes a tri-functional sulfate-activating complex. *Microbiology,* 150**,** 1681-6.

POL, A., DEN CAMP, H. J. O., MEES, S. G., KERSTEN, M. A. & VAN DER DRIFT, C. 1994. Isolation of a dimethylsulfide-utilizing *Hyphomicrobium* species and its application in biofiltration of polluted air. *Biodegradation,* 5**,** 105-112.

POL, A., RENKEMA, G. H., TANGERMAN, A., WINKEL, E. G., ENGELKE, U. F., DE BROUWER, A. P., LLOYD, K. C., ARAIZA, R. S., VAN DEN HEUVEL, L. & OMRAN, H. 2018. Mutations in SELENBP1, encoding a novel human methanethiol oxidase, cause extraoral halitosis. *Nature Genetics,* 50**,** 120.

QUINLAN, A. R. & HALL, I. M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics,* 26**,** 841-842.

QUINN, P. & BATES, T. 2011. The case against climate regulation via oceanic phytoplankton sulphur emissions. *Nature,* 480**,** 51.

RISSMAN, A. I., MAU, B., BIEHL, B. S., DARLING, A. E., GLASNER, J. D. & PERNA, N. T. 2009. Reordering contigs of draft genomes using the Mauve aligner. *Bioinformatics,* 25**,** 2071-2073.

RUOKOJÄRVI, A., AATAMILA, M., HARTIKAINEN, T., OLKKONEN, M., SALMI, J., RUUSKANEN, J. & MARTIKAINEN, P. 2000. Removal of dimethyl sulphide from off-gas mixtures containing hydrogen sulphide and methanethiol by a biotrickling filter. *Environmental Technology,* 21**,** 1173-1180.

SAMBASIVARAO, D., SCRABA, D. G., TRIEBER, C. & WEINER, J. H. 1990. Organization of dimethyl sulfoxide reductase in the plasma membrane of *Escherichia coli*. *Journal of Bacteriology,* 172**,** 5938-5948.

SANTOS, A. A., VENCESLAU, S. S., GREIN, F., LEAVITT, W. D., DAHL, C., JOHNSTON, D. T. & PEREIRA, I. A. 2015. A protein trisulfide couples dissimilatory sulfate reduction to energy conservation. *Science,* 350**,** 1541-1545.

SCHÄFER, H., MYRONOVA, N. & BODEN, R. 2009. Microbial degradation of dimethylsulphide and related C1-sulphur compounds: organisms and pathways controlling fluxes of sulphur in the biosphere. *Journal of Experimental Botany,* 61**,** 315-334.

SCHMIDT, S., CHRISTEN, P., KIEFER, P. & VORHOLT, J. A. 2010. Functional investigation of methanol dehydrogenase-like protein XoxF in *Methylobacterium extorquens* AM1. *Microbiology,* 156**,** 2575-2586.

SEKOWSKA, A., KUNG, H.-F. & DANCHIN, A. 2000. Sulfur metabolism in *Escherichia coli* and related bacteria: facts and fiction. *Journal of Molecular Microbiology and Biotechnology,* 2**,** 145-177.

SEYMOUR, J. R., SIMO, R., AHMED, T. & STOCKER, R. 2010. Chemoattraction to dimethylsulfoniopropionate throughout the marine microbial food web. *Science,* 329**,** 342-5.

SIVELÄ, S. & SUNDMAN, V. 1975. Demonstration of *Thiobacillus*-type bacteria, which utilize methyl sulphides. *Archives of Microbiology,* 103**,** 303-304.

STEINER, E., BÖTH, D., LÖSSL, P., VILAPLANA, F., SCHNELL, R. & SCHNEIDER, G. 2014. CysK2 from *Mycobacterium tuberculosis* is an O-phospho-l-serine-dependent S-sulfocysteine synthase. *Journal of Bacteriology*, 196, 3410–3420.

STENSMYR, M. C., URRU, I., COLLU, I., CELANDER, M., HANSSON, B. S. & ANGIOY, A.-M. 2002. Pollination: Rotting smell of dead-horse arum florets. *Nature,* 420**,** 625.

STEUDLER, P. & PETERSON, B. 1984. Contribution of gaseous sulphur from salt marshes to the global sulphur cycle. *Nature,* 311**,** 455.

STUDER, A., STUPPERICH, E., VUILLEUMIER, S. & LEISINGER, T. 2001. Chloromethane: tetrahydrofolate methyl transfer by two proteins from *Methylobacterium chloromethanicum* strain CM4. *European Journal of Biochemistry,* 268**,** 2931-2938.

SUN, Y., WANG, M., LI, L., ZHOU, L., WANG, X., ZHENG, P., YU, H., LI, C. & SUN, S. 2017. Molecular identification of methane monooxygenase and quantitative analysis of methanotrophic endosymbionts under laboratory maintenance in *Bathymodiolus platifrons* from the South China Sea. *PeerJ,* 5**,** e3565.

SUNDA, W., KIEBER, D. J., KIENE, R. P. & HUNTSMAN, S. 2002. An antioxidant function for DMSP and DMS in marine algae. *Nature,* 418**,** 317-20.

SUYLEN, G. & KUENEN, J. 1986. Chemostat enrichment and isolation of *Hyphomicrobium* EG. *Antonie van Leeuwenhoek,* 52**,** 281-293.

SUYLEN, G., STEFESS, G. & KUENEN, J. 1986. Chemolithotrophic potential of a *Hyphomicrobium* species, capable of growth on methylated sulphur compounds. *Archives of Microbiology,* 146**,** 192-198.

TALLANT, T. C. & KRZYCKI, J. A. 1997. Methylthiol: coenzyme M methyltransferase from *Methanosarcina barkeri*, an enzyme of methanogenesis from dimethylsulfide and methylmercaptopropionate. *Journal of Bacteriology,* 179**,** 6902-6911.

TALLANT, T. C., PAUL, L. & KRZYCKI, J. A. 2001. The MtsA subunit of the methylthiol: coenzyme M methyltransferase of *Methanosarcina barkeri* catalyses both half-reactions of corrinoid-dependent dimethylsulfide: coenzyme M methyl transfer. *Journal of Biological Chemistry,* 276**,** 4485-4493.

TANGERMAN, A. & WINKEL, E. G. 2013. Volatile sulfur compounds as the cause of bad breath: a review. *Phosphorus, Sulfur, and Silicon and the Related Elements,* 188**,** 396-402.

THOMPSON, A. S., OWENS, N. & MURRELL, J. C. 1995. Isolation and characterization of methanesulfonic acid-degrading bacteria from the marine environment. *Applied and Environmental Microbiology,* 61**,** 2388-2393.

TISHKOV, V. I., GALKIN, A. G., MARCHENKO, G., EGOROVA, O. A., SHELUHO, D. V., KULAKOVA, L. B., DEMENTIEVA, L. & EGOROV, A. 1993. Catalytic properties and stability of a *Pseudomonas* sp. 101 formate dehydrogenase mutants containing Cys-255-Ser and Cys-255-Met replacements. *Biochemical and Biophysical Research Communications,* 192**,** 976-981.

TYANOVA, S., TEMU, T. & COX, J. 2016a. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature protocols,* 11**,** 2301.

TYANOVA, S., TEMU, T., SINITCYN, P., CARLSON, A., HEIN, M. Y., GEIGER, T., MANN, M. & COX, J. 2016b. The Perseus computational platform for comprehensive analysis of (prote) omics data. *Nature Methods,* 13**,** 731.

VAN DER PLOEG, J. R., CUMMINGS, N. J., LEISINGER, T. & CONNERTON, I. F. 1998. *Bacillus subtilis* genes for the utilization of sulfur from aliphatic sulfonates. *Microbiology,* 144**,** 2555-61.

VAN DER PLOEG, J. R., IWANICKA-NOWICKA, R., BYKOWSKI, T., HRYNIEWICZ, M. M. & LEISINGER, T. 1999. The *Escherichia coli ssuEADCB* gene cluster is

required for the utilization of sulfur from aliphatic sulfonates and is regulated by the transcriptional activator Cbl. *Journal of Biological Chemistry,* 274**,** 29358-29365.

VERMEIJ, P. & KERTESZ, M. A. 1999. Pathways of Assimilative Sulfur Metabolism in *Pseudomonas putida*. *Journal of Bacteriology,* 181**,** 5833-5837.

VORHOLT, J. A. 2002. Cofactor-dependent pathways of formaldehyde oxidation in methylotrophic bacteria. *Archives of Microbiology,* 178**,** 239-249.

VORHOLT, J. A., MARX, C. J., LIDSTROM, M. E. & THAUER, R. K. 2000. Novel formaldehyde-activating enzyme in *Methylobacterium extorquens* AM1 required for growth on methanol. *Journal of Bacteriology,* 182**,** 6645-6650.

WANG, R. 2012. Physiological implications of hydrogen sulfide: a whiff exploration that blossomed. *Physiological Reviews,* 92**,** 791-896.

WANG, S., MALTRUD, M., ELLIOTT, S., CAMERON-SMITH, P. & JONKO, A. 2018. Influence of dimethyl sulfide on the carbon cycle and biological production. *Biogeochemistry,* 138**,** 49-68.

WATTS, S. F. 2000. The mass budgets of carbonyl sulfide, dimethyl sulfide, carbon disulfide and hydrogen sulfide. *Atmospheric Environment,* 34**,** 761-779.

WICHT, D. K. 2016. The reduced flavin-dependent monooxygenase SfnG converts dimethylsulfone to methanesulfinate. *Archives of Biochemistry and Biophysics,* 604**,** 159-166.

XIA, Y., LÜ, C., HOU, N., XIN, Y., LIU, J., LIU, H. & XUN, L. 2017. Sulfide production and oxidation by heterotrophic bacteria under aerobic conditions. *The ISME Journal,* 11**,** 2754.

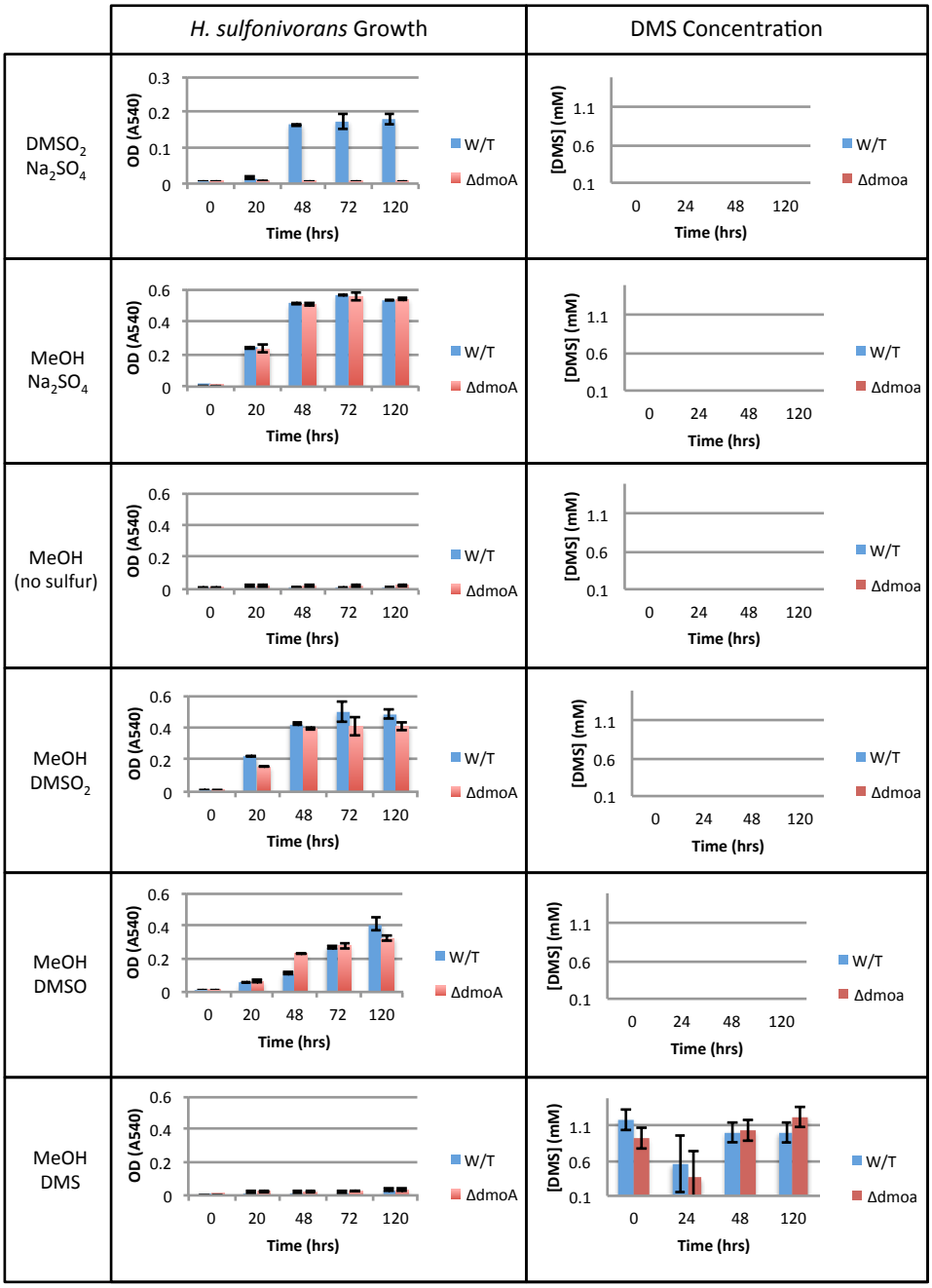
YE, J., COULOURIS, G., ZARETSKAYA, I., CUTCUTACHE, I., ROZEN, S. & MADDEN, T. L. 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics,* 13**,** 134.

YI, Z., ZHENG, L., WU, T. & WANG, X. 2013. Contribution of aboveground plants, the rhizosphere and root-free-soils to total COS and DMS fluxes at three key growth stages in rice paddies. *Agriculture, Ecosystems & Environment,* 179**,** 11-17.

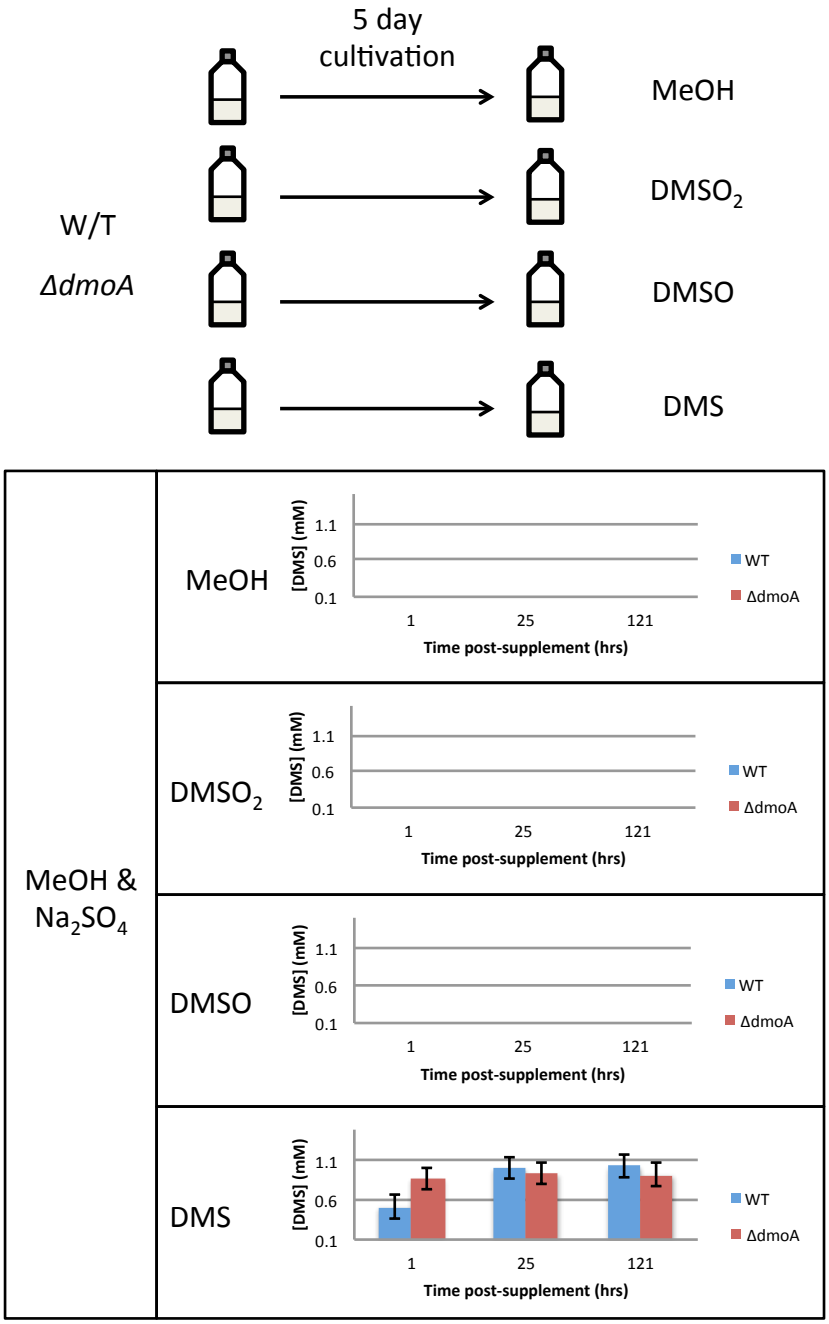
YOCH, D. C. 2002. Dimethylsulfoniopropionate: its sources, role in the marine food web, and biological degradation to dimethylsulfide. *Applied and Environmental Microbiology,* 68**,** 5804-5815.

YOSHIDA, Y., NEGISHI, M. & NAKANO, Y. 2003. Homocysteine biosynthesis pathways of *Streptococcus anginosus*. *FEMS Microbiology Letters,* 221**,** 277-284.

Appendixes



Supplementary Figure S3.2-1: *Phenotyping of the* H. sulfonivorans *S1 strain.* The *H. sulfonivorans* S1 WT and *ΔdmoA* strains were cultivated in triplicate on a respective sole carbon/sulfur source of $DMSO_2$ (20mM)/Na_2SO_4 (1mM), MeOH (0.5%)/Na_2SO_4 (1mM), MeOH (0.5%)/no sulfur, MeOH (0.5%)/$DMSO_2$ (1mM), MeOH (0.5%)/DMSO (1mM), or MeOH (0.5%)/DMS (1mM). The growth of each strain was then assessed over the next 120 h by measuring culture OD_{540}, while culture headspace DMS concentration was measured using gas chromatography.



Supplementary Figure S3.2-2: *MSC supplementation of turbid* H. sulfonivorans *cultures. H. sulfonivorans* W/T and *Δ*dmoA strains were cultivated for 120 h on a sole carbon source of MeOH (20mM). They were then supplemented in triplicate with either 1mM MeOH (control), $DMSO_2$, DMSO or DMS, and DMS concentration within each culture measured over time using gas chromatography.

Supplementary Table S3.2-1: *Phenotyping of* H. sulfonivorans *S1 WT and Δ*dmoA *strains for the utilization of MSCs as a sole carbon source.* P-values have been calculated via Student's T-test from triplicate OD_{540} samples at T_{144}.

H. sulfonivorans fed-batch phenotyping cultures (carbon source) at T_{144}									
carbon source	WT				*ΔdmoA*				p-value (WT vs. *ΔdmoA*)
	mean OD_{540}	relative growth yield	p-value vs. control	phenotype	mean OD_{540}	relative growth yield	p-value vs. control	phenotype	
MeOH	0.090	5.87	0.000	++	0.091	4.95	0.000	++	0.740
no carbon	0.015	1.00	1.000	-	0.018	1.00	1.000	-	0.257
$DMSO_2$	0.033	2.15	0.002	+	0.018	1.00	1.000	-	0.004
DMSO	0.016	1.02	0.882	-	0.015	0.82	0.130	-	0.644
DMS	0.017	1.09	0.630	-	0.018	1.00	1.000	-	0.461
MSA	0.017	1.09	0.508	-	0.016	0.87	0.020	-	0.635

Supplementary Table S3.2-2: *Phenotyping of* H. sulfonivorans *S1 WT and Δ*dmoA *strains for the utilization of MSCs as a sole sulfur source.* P-values have been calculated via Student's T-test from triplicate OD_{540} samples at T_{50}.

H. sulfonivorans batch phenotyping cultures (sulfur source) at T_{50}									
sulfur source	WT				*ΔdmoA*				p-value (WT vs. *ΔdmoA*)
	mean OD_{540}	relative growth yield	p-value vs. control	phenotype	mean OD_{540}	relative growth yield	p-value vs. control	phenotype	
MeOH	0.329	31.81	0.010	++	0.283	17.67	0.004	+	0.172
no carbon	0.010	1.00	1.000	-	0.016	1.00	1.000	-	0.136
$DMSO_2$	0.244	23.58	0.015	++	0.308	19.25	0.002	++	0.070
DMSO	0.365	35.32	0.002	++	0.370	23.13	0.007	++	0.831
DMS	0.059	5.68	0.000	+	0.061	3.79	0.009	+	0.650
MSA	0.291	28.13	0.000	++	0.284	17.73	0.002	+	0.622

```
Draft   1                                   TCAATATGAGAGTTTGATCCTGGCTCAGAACGAAC  35

Draft  36    GCTGGCGGCAGGCCTAACACATGCAAGTCGAACGCCCCGCAAGGGGAGTGGCAGACGGGT  95
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref  1    GCTGGCGGCAGGCCTAACACATGCAAGTCGAACGCCCCGCAAGGGGAGTGGCAGACGGGT  60

Draft  96    GAGTAACGCGTGGGAACCTTCCCAATGGTGCGGAATAGCCCAGGGAAACTTGGAGTAATA  155
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref 61    GAGTAACGCGTGGGAACCTTCCCAATGGTGCGGAATAGCCCAGGGAAACTTGGAGTAATA  120

Draft  156   CCGCATAAGCCCTTAGGGGGAAAGATTTATCGCCATTGGATGGGCCCGCGTCGGATTAGC  215
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref 121   CCGCATAAGCCCTTAGGGGGAAAGATTTATCGCCATTGGATGGGCCCGCGTCGGATTAGC  180

Draft  216   TAGTTGGTAGGGTAACGGCCTACCAAGGCGACGATCCGTAGCTGGTTTGAGAGAATGACC  275
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref 181   TAGTTGGTAGGGTAACGGCCTACCAAGGCGACGATCCGTAGCTGGTTTGAGAGAATGACC  240

Draft  276   AGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTT  335
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref 241   AGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATCTT  300

Draft  336   GGACAATGGGCGCAAGCCTGATCCAGCCATGCCGCGTGAGTGACGAAGGCCTTAGGGTTG  395
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref 301   GGACAATGGGCGCAAGCCTGATCCAGCCATGCCGCGTGAGTGACGAAGGCCTTAGGGTTG  360

Draft  396   TAAAGCTCTTTTGCCGGGGACGATAATGACGGTACCCGGAGAATAAGTCCCGGCTAACTT  455
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref 361   TAAAGCTCTTTTGCCGGGGACGATAATGACGGTACCCGGAGAATAAGTCCCGGCTAACTT  420

Draft  456   CGTGCCAGCAGCCGCGGTAATACGAAGGGGACTAGCGTTGTTCGGAATTACTGGGCGTAA  515
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref 421   CGTGCCAGCAGCCGCGGTAATACGAAGGGGACTAGCGTTGTTCGGAATTACTGGGCGTAA  480

Draft  516   AGCGCACGTAGGCGGATTATTAAGTCAGGGGTGAAATCCCGGGGCTCAACCTCGGAACTG  575
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref 481   AGCGCACGTAGGCGGATTATTAAGTCAGGGGTGAAATCCCGGGGCTCAACCTCGGAACTG  540

Draft  576   CCTTTGATACTGGTAATCTAGAGTCCGATAGAGGTGGGTGGAATTCCTAGTGTAGAGGTG  635
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref 541   CCTTTGATACTGGTAATCTAGAGTCCGATAGAGGTGGGTGGAATTCCTAGTGTAGAGGTG  600

Draft  636   AAATTCGTAGATATTAGGAAGAACACCAGTGGCGAAGGCGGCCCACTGGATCGGTACTGA  695
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
S1-ref 601   AAATTCGTAGATATTACGAAGAACACCAGTGGCGAAGGCGGCCCACTGGATCGGTACTGA  660

Draft  696   CGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGT  755
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref 661   CGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGT  720
```

Supplementary Figure S3.3-1a: *BLASTn results of the 16S RNA gene sequence from* H. sulfonivorans *S1 2018 draft genome versus the* H. sulfonivorans *S1 partial 16S RNA gene sequence (S1-ref, NCBI Accession NR_025082.1) submitted by Borodina et al. (2000).*

```
Draft   756   AAACGATGGATGCTAGCCGTCGGCAAGCTTGCTTGTCGGTGGCGCAGCTAACGCATTAAG  815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref  721   AAACGATGGATGCTAGCCGTCGGCAAGCTTGCTTGTCGGTGGCGCAGCTAACGCATTAAG  780

Draft   816   CATCCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGC  875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref  781   CATCCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGC  840

Draft   876   ACAAGCGGTGGAGCATGTGGTTTAATTCGACGCAACGCGAAGAACCTTACCAGCTCTTGA  935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref  841   ACAAGCGGTGGAGCATGTGGTTTAATTCGACGCAACGCGAAGAACCTTACCAGCTCTTGA  900

Draft   936   CATGGTCGGTCGGTTTCCAGAGATGGATTCCTCCTAGCAATAGGCCGATACACAGGTGCT  995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref  901   CATGGTCGGTCGGTTTCCAGAGATGGATTCCTCCTAGCAATAGGCCGATACACAGGTGCT  960

Draft   996   GCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAAC  1055
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref  961   GCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAAC  1020

Draft   1056  CCTCGCCATTAGTTGCCATCATTAAGTTGGGCACTCTAGTGGGACTGCCGGTGATAAGCC  1115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref  1021  CCTCGCCATTAGTTGCCATCATTAAGTTGGGCACTCTAGTGGGACTGCCGGTGATAAGCC  1080

Draft   1116  GGAGGAAGGTGGGGATGACGTCAAGTCATCATGGCCCTTACGGGCTGGGCTACACACGTG  1175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref  1081  GGAGGAAGGTGGGGATGACGTCAAGTCATCATGGCCCTTACGGGCTGGGCTACACACGTG  1140

Draft   1176  CTACAATGGCGGTGACAGTGGGCAGCAACACAGCGATGTGAAGCTAATCTCAAAAAGCCG  1235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref  1141  CTACAATGGCGGTGACAGTGGGCAGCAACACAGCGATGTGAAGCTAATCTCAAAAAGCCG  1200

Draft   1236  TCTCAGTTCGGATTGGTCTCTGCAACTCGAGACCATGAAGTTGGAATCGCTAGTAATCGC  1295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref  1201  TCTCAGTTCGGATTGGTCTCTGCAACTCGAGACCATGAAGTTGGAATCGCTAGTAATCGC  1260

Draft   1296  GGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCAT  1355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref  1261  GGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCAT  1320

Draft   1356  GGGAGTTGGTTCTACCCGAAGGCGATGCGCTAACCCGCAAGGGAGGCAGTCGACCACGGT  1415
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
S1-ref  1321  GGGAGTTGGTTCTACCCGAAGGCGATGCGCTAACCCGCAAGGGAGGCAGTCGACCACGGT  1380

Draft   1416  AGGGTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCTGGAT  1475
              ||||||||||||||||||||||||||||||||
S1-ref  1381  AGGGTCAGCGACTGGGGTGAAGTCGTAACAA                               1411

Draft   1476  CACCTCCTTTCT                                                  1487
```

Supplementary Figure S3.3-1b: *BLASTn results of the 16S RNA gene sequence from* H. sulfonivorans *S1 draft 2018 (Draft) versus the* H. sulfonivorans *S1 partial 16S RNA gene sequence (S1-ref, NCBI Accession NR_025082.1) submitted by Borodina et al. (2000).*

Supplementary Table S4.2-1a: *Comparative transcriptomics of* H. sulfonivorans*, genes of interest count data. Unidentified proteins marked as N/A.*

Gene ID	Gene name (or product)	Contig	Gene locus	Metabolic Pathway	Annotation	
					KO number	NCBI WGS
C6Y62_00305	*pdo*	1	63122:64012 Forward	S		MBL fold metallo-hydrolase
C6Y62_00310	*rhd-sqrB*	1	64046:65710 Forward	S	K17218	TIGR01244 family phosphatase
C6Y62_00370	*mxaI*	1	75675:75962 Reverse	MeOH	K14029	methanol dehydrogenase
C6Y62_00375	*mxaG*	1	76050:76634 Reverse	MeOH	K16255	cytochrome c(L)%2C periplasmic
C6Y62_00380	*mxaJ*	1	76631:77533 Reverse	MeOH		methanol oxidation system protein MoxJ
C6Y62_00385	*mxaF*	1	77733:79622 Reverse	MeOH	K14028	PQQ-dependent dehydrogenase, methanol/ethanol family
C6Y62_00835	*msuD*	1	183693:184880 Forward	MSC	K04091	alkanesulfonate monooxygenase, FMNH(2)-dependent
C6Y62_00880	*sfnF*	1	193988:194572 Forward	MSC	K00299	FMN reductase
C6Y62_00885	*sfnG*	1	194603:195718 Forward	MSC	K17228	dimethyl sulfone monooxygenase SfnG
C6Y62_02365	*fdhA*	1	526913:527914 Forward	HCHO	K12957	alcohol dehydrogenase
C6Y62_02980	*fdh*	1	650742:651947 Forward	HCHO	K00122	NAD-dependent formate dehydrogenase
C6Y62_04625	*acs*	1	996883:998838 Forward	S	K01895	acetate--CoA ligase
C6Y62_04710	*folD*	1	1014098:1014976 Forward	HCHO	K01491	bifunctional methylene H4F dehydrogenase/methenyl H4F cyclohydrolase FolD
C6Y62_04850	*fhs*	1	1047741:1049417 Forward	HCHO	K01938	formate--tetrahydrofolate ligase
C6Y62_05070	*metZ*	1	1099534:1100751 Reverse	S	K10764	O-succinylhomoserine sulfhydrylase
C6Y62_05220	*fae*	2	15756:16307 Forward	HCHO		aldehyde-activating protein
C6Y62_05595	*fdh*	2	99340:100545 Reverse	HCHO	K00122	NAD-dependent formate dehydrogenase
C6Y62_08560	*cysK*	4	50688:51728 Reverse	S	K01738	cysteine synthase A
C6Y62_09260	(catalase)	4	214166:215644 Forward	MSC	K03781	catalase
C6Y62_09280	*cysE*	4	218851:219672 Forward	S	K00640	serine O-acetyltransferase
C6Y62_10250	*fdhA*	4	421692:422732 Reverse	HCHO	K18369	alcohol dehydrogenase
C6Y62_10660	*soxY*	4	530308:530778 Forward	S	K17226	thiosulfate oxidation carrier protein SoxY
C6Y62_10665	*soxZ*	4	530792:531115 Forward	S	K17226	thiosulfate oxidation carrier complex protein SoxZ

Supplementary Table S4.2-1b: *Comparative transcriptomics of* H. sulfonivorans, *genes of interest count data (continued). Unidentified proteins marked as N/A.*

Gene ID	Gene name (or product)	Contig	Gene locus	Metabolic Pathway	Annotation	
					KO number	NCBI WGS
C6Y62_12280	*soxYZ*	5	181795:182595 Reverse	S	K17226	quinoprotein dehydrogenase-associated SoxYZ-like carrier
C6Y62_12500	*fchC*	5	230487:231296 Reverse	HCHO	K00202	formylmethanofuran dehydrogenase subunit C
C6Y62_12505	*fchD*	5	231293:232192 Reverse	HCHO	K00672	formylmethanofuran--tetrahydromethanopterin N-formyltransferase
C6Y62_12510	*fchA*	5	232189:233856 Reverse	HCHO	K00200	formylmethanofuran dehydrogenase subunit A
C6Y62_12515	*fchB*	5	233875:235236 Reverse	HCHO	K00201	formylmethanofuran dehydrogenase
C6Y62_12520	*xoxF*	5	235928:237925 Forward	MeOH		PQQ-dependent dehydrogenase%2C methanol/ethanol family
C6Y62_12595	*fae*	5	249305:249817 Reverse	HCHO	K10713	formaldehyde-activating enzyme
C6Y62_12610	*mch*	5	251956:252921 Reverse	HCHO	K01499	methenyltetrahydromethanopterin cyclohydrolase
C6Y62_12620	*mtdB*	5	254089:254988 Reverse	HCHO	K10714	methylenetetrahydromethanopterin dehydrogenase
C6Y62_12675	*cysJ*	5	263720:265498 Forward	S	K00380	sulfite reductase subunit alpha
C6Y62_12695	*cysNC*	5	269438:270901 Reverse	S	K00955	sulfate adenylyltransferase
C6Y62_12700	*cysD*	5	270901:271806 Reverse	S	K00957	sulfate adenylyltransferase subunit CysD
C6Y62_12705	*cysH*	5	271803:272618 Reverse	S	K00390	phosphoadenylyl-sulfate reductase
C6Y62_13025	*fae*	5	343050:343589 Forward	HCHO		formaldehyde-activating enzyme
C6Y62_13185	*sfnF*	5	385211:385756 Forward	MSC	K00299	FMN reductase
C6Y62_13190	*sfnG*	5	385844:386962 Forward	MSC	K17228	dimethyl sulfone monooxygenase SfnG
C6Y62_13200	*dmoB*	5	388235:388765 Forward	MSC		flavin reductase
C6Y62_13210	*dmoA*	5	390172:391614 Forward	MSC	K20938	5%2C10-methylene tetrahydromethanopterin reductase
C6Y62_13220	*soxC*	5	392325:393596 Forward	S	K17225	sulfite dehydrogenase
C6Y62_13225	*soxD*	5	393589:394167 Forward	S	K22622	cytochrome C
C6Y62_14215	*soxYZ*	6	179225:180067 Reverse	S	K17226	quinoprotein dehydrogenase-associated SoxYZ-like carrier
C6Y62_15095	*cysC*	7	3722:5146 Reverse	S	K00955	adenylyl-sulfate kinase
C6Y62_15390	catalase	7	69793:70755 Forward	MSC	K03781	catalase
C6Y62_15815	*cysK*	8	51189:52190 Reverse	S	K01738	cysteine synthase A

Supplementary Table S4.2-1c: *Comparative transcriptomics of* H. sulfonivorans, *genes of interest count data (continued). Unidentified proteins marked as N/A.*

Gene ID	Gene name (or product)	Metabolic Pathway	Read Count						MSC versus Control		
			Control			MSC			p-value	padj	Fold change
			1	2	3	1	2	3			
C6Y62_00305	*pdo*	S	1,258	1,144	1,200	1,004	844	1,421	0.127	0.245	0.82
C6Y62_00310	*rhd-sqrB*	S	4,488	5,404	4,933	5,616	4,030	6,036	0.623	0.766	0.97
C6Y62_00370	*mxaI*	MeOH	1,569,426	1,562,702	1,339,589	740,536	490,536	711,088	0.000	0.000	0.40
C6Y62_00375	*mxaG*	MeOH	77,800	97,297	102,738	79,267	48,811	69,308	0.001	0.003	0.65
C6Y62_00380	*mxaJ*	MeOH	43,264	54,920	56,983	52,793	35,352	47,352	0.071	0.154	0.80
C6Y62_00385	*mxaF*	MeOH	8,466,857	9,372,611	8,486,117	3,868,527	2,569,690	3,672,817	0.000	0.000	0.35
C6Y62_00835	*msuD*	MSC	87	113	141	102	41	155	0.270	0.428	0.72
C6Y62_00880	*sfnF*	MSC	139	177	137	189	133	318	0.362	0.531	1.23
C6Y62_00885	*sfnG*	MSC	268	268	202	251	169	411	0.941	0.970	0.98
C6Y62_02365	*fdhA*	HCHO	3,711	4,002	4,035	4,401	3,198	5,328	0.957	0.978	1.00
C6Y62_02980	*fdh*	HCHO	12,346	6,702	7,056	2,752	2,098	3,631	0.000	0.000	0.31
C6Y62_04625	*acs*	S	33,130	34,654	28,774	33,344	25,618	39,812	0.363	0.533	0.93
C6Y62_04710	*folD*	HCHO	15,955	19,782	18,644	16,884	12,038	23,878	0.035	0.089	0.87
C6Y62_04850	*fhs*	HCHO	37,224	47,869	45,193	48,608	33,621	62,793	0.967	0.983	1.00
C6Y62_05070	*metZ*	S	6,356	8,382	7,158	6,594	3,633	8,276	0.014	0.040	0.74
C6Y62_05220	*fae*	HCHO	1,062	1,406	1,411	3,083	1,819	3,372	0.000	0.000	1.92
C6Y62_05595	*fdh*	HCHO	563	720	561	774	596	960	0.184	0.322	1.16
C6Y62_08560	*cysK*	S	2,199	2,402	2,349	3,465	2,658	3,896	0.001	0.003	1.32
C6Y62_09260	(catalase)	MSC	8,587	12,496	11,793	131,053	94,803	139,529	0.000	0.000	10.27
C6Y62_09280	*cysE*	S	3,570	4,420	4,027	5,551	3,602	6,333	0.046	0.110	1.16
C6Y62_10250	*fdhA*	HCHO	14,245	6,871	7,408	2,754	2,018	4,163	0.000	0.000	0.29
C6Y62_10660	*soxY*	S	2,015	1,858	1,568	1,535	1,041	2,102	0.046	0.110	0.77
C6Y62_10665	*soxZ*	S	466	415	313	335	315	406	0.313	0.474	0.82

Supplementary Table S4.2-1d: *Comparative transcriptomics of* H. sulfonivorans, *genes of interest count data (continued). Unidentified proteins marked as N/A.*

Gene ID	Gene name (or product)	Metabolic Pathway	Read Count						MSC versus Control		
			Control			MSC			p-value	padj	Fold change
			1	2	3	1	2	3			
C6Y62_12280	*soxYZ*	S	370	589	639	640	511	773	0.452	0.617	1.13
C6Y62_12500	*fchC*	HCHO	7,862	8,351	7,538	10,787	7,699	13,832	0.005	0.017	1.22
C6Y62_12505	*fchD*	HCHO	7,963	9,293	9,103	11,746	8,756	14,947	0.000	0.001	1.22
C6Y62_12510	*fchA*	HCHO	7,111	9,375	10,459	18,062	13,208	22,240	0.000	0.000	1.82
C6Y62_12515	*fchB*	HCHO	28,870	36,901	34,029	51,405	33,924	57,967	0.000	0.002	1.30
C6Y62_12520	*xoxF*	MeOH	8,680	7,019	6,202	6,036	4,096	7,299	0.004	0.016	0.72
C6Y62_12595	*fae*	HCHO	107,676	176,642	182,583	417,029	272,697	468,352	0.000	0.000	2.30
C6Y62_12610	*mch*	HCHO	36	45	48	17	41	16	0.175	0.313	0.49
C6Y62_12620	*mtdB*	HCHO	21,886	28,441	26,762	27,445	19,394	41,818	0.818	0.900	1.02
C6Y62_12675	*cysJ*	S	8,269	9,800	11,182	10,741	17,686	13,390	0.203	0.348	1.30
C6Y62_12695	*cysNC*	S	15,432	17,805	22,360	13,844	1,851	18,424	0.060	0.135	0.40
C6Y62_12700	*cysD*	S	21	22	24	31	8,371	26	0.380	0.551	2.98
C6Y62_12705	*cysH*	S	18	15	18	20	2,512	16	0.476	0.640	2.33
C6Y62_13025	*fae*	HCHO	218	382	495	1,017	759	1,409	0.000	0.000	2.74
C6Y62_13185	*sfnF*	MSC	4,347	3,159	3,045	128,081	81,253	140,580	0.000	0.000	30.25
C6Y62_13190	*sfnG*	MSC	2,486	2,505	2,635	104,544	72,850	108,053	0.000	0.000	34.11
C6Y62_13200	*dmoB*	MSC	522	559	646	39,267	27,112	35,220	0.000	0.000	53.98
C6Y62_13210	*dmoA*	MSC	2,334	1,851	1,555	1,766,014	1,279,246	2,038,556	0.000	0.000	815.81
C6Y62_13220	*soxC*	S	144	123	104	55,172	37,820	61,873	0.000	0.000	379.97
C6Y62_13225	*soxD*	S	252	226	242	25,887	17,941	27,826	0.000	0.000	90.47
C6Y62_14215	*soxYZ*	S	7,217	9,636	8,572	32,486	24,852	39,177	0.000	0.000	3.48
C6Y62_15095	*cysC*	S	4,821	6,376	5,525	7,518	5,433	7,430	0.187	0.326	1.12
C6Y62_15390	catalase	MSC	2,877	4,058	4,531	7,298	4,906	6,904	0.003	0.011	1.55
C6Y62_15815	*cysK*	S	6,498	6,861	7,550	7,948	8,752	9,858	0.271	0.428	1.17

Supplementary Table S4.3-2aa: *Comparative proteomics of* H. sulfonivorans*, Experiment 2, proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			Unique Peptides									MS/MS count								
Protein ID	Protein Name	Pathway	Control			Sulfur			Carbon/Sulfur			Control			Sulfur			Carbon/Sulfur		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
C6Y62_00305	Pdo	S	2	2	3	3	4	4	2	1	2	0	0	0	1	2	2	1	0	0
C6Y62_00310	Rhd-SqrB	S	17	17	14	14	14	12	13	11	13	18	14	12	13	12	7	11	2	10
C6Y62_00370	MxaI	MeOH	4	5	5	6	5	5	3	1	1	3	2	5	6	3	4	1	0	0
C6Y62_00375	MxaG	MeOH	12	12	13	14	13	14	11	6	10	28	21	29	26	22	28	6	1	11
C6Y62_00380	MxaJ	MeOH	24	24	23	22	20	20	20	19	18	42	43	39	47	36	39	19	7	25
C6Y62_00385	MxaF	MeOH	62	64	65	64	63	62	57	53	60	635	661	840	696	706	722	289	212	437
C6Y62_00835	MsuD	MSC	2	5	3	29	30	29	27	25	28	1	1	1	58	60	57	37	16	37
C6Y62_00880	SfnF	MSC	0	0	0	9	9	9	6	5	5	0	0	0	16	18	14	4	4	3
C6Y62_00885	SfnG	MSC	8	8	11	28	30	30	25	22	21	3	4	7	73	75	80	36	17	32
C6Y62_02365	FdhA	HCHO	13	14	13	15	14	14	13	12	14	10	12	22	18	15	16	18	10	23
C6Y62_02980	Fdh	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_04625	Acs	S	30	29	32	33	32	33	33	32	34	32	27	33	29	32	29	46	24	50
C6Y62_04710	FolD	HCHO	25	26	27	26	26	25	25	24	25	73	82	84	84	75	75	84	48	75
C6Y62_04850	Fhs	HCHO	48	49	51	50	48	48	47	43	48	119	109	134	123	124	117	107	58	111
C6Y62_05070	MetZ	S	23	22	22	24	23	26	25	21	24	25	24	27	25	28	29	29	13	29
C6Y62_05220	Fae	HCHO	9	10	11	11	10	10	9	9	9	13	18	23	19	18	19	16	7	12
C6Y62_05595	Fdh	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_08560	CysK	S	10	9	10	9	10	11	12	8	7	7	9	8	5	9	9	13	3	7
C6Y62_09260	catalase	MSC	21	26	25	28	28	27	32	31	32	27	20	20	29	30	31	55	22	61
C6Y62_09280	CysE	S	14	14	13	14	12	12	14	15	14	16	13	15	15	14	14	13	8	13
C6Y62_10250	FdhA	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_10660	SoxY	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_10665	SoxZ	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Supplementary Table S4.3-2ab: *Comparative proteomics of* H. sulfonivorans*, Experiment 2, proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			Intensity								
Protein ID	Protein Name	Pathway	Control			Sulfur			Carbon/Sulfur		
			1	2	3	1	2	3	1	2	3
C6Y62_00305	Pdo	S	2.81E+06	3.88E+06	6.12E+06	6.75E+06	1.05E+07	1.01E+07	2.92E+06	3.91E+04	4.49E+06
C6Y62_00310	Rhd-SqrB	S	2.45E+08	2.46E+08	2.49E+08	1.50E+08	1.78E+08	1.30E+08	1.52E+08	4.25E+06	2.05E+08
C6Y62_00370	MxaI	MeOH	1.37E+07	1.94E+07	4.57E+07	4.43E+07	2.98E+07	4.35E+07	4.05E+06	1.43E+05	1.81E+06
C6Y62_00375	MxaG	MeOH	6.58E+08	4.68E+08	9.08E+08	6.37E+08	6.34E+08	5.86E+08	5.97E+07	9.52E+05	2.06E+08
C6Y62_00380	MxaJ	MeOH	2.64E+09	2.31E+09	2.39E+09	2.26E+09	2.00E+09	1.80E+09	3.38E+08	5.91E+06	1.12E+09
C6Y62_00385	MxaF	MeOH	5.91E+10	4.91E+10	6.72E+10	4.37E+10	4.49E+10	4.86E+10	1.73E+10	6.40E+08	4.61E+10
C6Y62_00835	MsuD	MSC	2.11E+06	3.82E+06	1.87E+06	2.98E+09	3.76E+09	3.07E+09	8.49E+08	1.98E+07	1.10E+09
C6Y62_00880	SfnF	MSC	0.00	0.00	0.00	6.49E+08	8.17E+08	7.48E+08	6.94E+07	1.33E+06	8.16E+07
C6Y62_00885	SfnG	MSC	2.31E+07	4.00E+07	6.52E+07	9.94E+09	9.86E+09	1.11E+10	1.59E+09	3.25E+07	1.80E+09
C6Y62_02365	FdhA	HCHO	2.26E+08	2.47E+08	2.26E+08	2.04E+08	2.36E+08	2.19E+08	3.48E+08	8.15E+06	5.03E+08
C6Y62_02980	Fdh	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_04625	Acs	S	9.99E+08	8.29E+08	1.03E+09	7.23E+08	7.25E+08	6.75E+08	1.35E+09	3.21E+07	1.81E+09
C6Y62_04710	FolD	HCHO	7.27E+09	7.40E+09	9.89E+09	6.79E+09	7.07E+09	6.10E+09	6.75E+09	1.86E+08	7.56E+09
C6Y62_04850	Fhs	HCHO	1.02E+10	8.33E+09	9.24E+09	7.35E+09	9.07E+09	7.16E+09	8.21E+09	1.69E+08	1.40E+10
C6Y62_05070	MetZ	S	6.34E+08	5.50E+08	6.60E+08	4.56E+08	4.87E+08	4.31E+08	3.43E+08	1.75E+07	6.62E+08
C6Y62_05220	Fae	HCHO	7.50E+08	9.09E+08	1.21E+09	1.00E+09	9.08E+08	8.94E+08	4.98E+08	6.01E+07	4.80E+08
C6Y62_05595	Fdh	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_08560	CysK	S	6.50E+07	4.56E+07	5.18E+07	4.98E+07	6.64E+07	5.99E+07	9.54E+07	5.46E+06	1.02E+08
C6Y62_09260	catalase	MSC	2.51E+08	3.14E+08	3.32E+08	4.64E+08	3.95E+08	4.07E+08	2.13E+09	3.79E+07	2.18E+09
C6Y62_09280	CysE	S	2.54E+08	2.15E+08	2.81E+08	2.04E+08	1.81E+08	1.86E+08	3.29E+08	7.79E+06	3.55E+08
C6Y62_10250	FdhA	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_10660	SoxY	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_10665	SoxZ	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Supplementary Table S4.3-2ac: *Comparative proteomics of* H. sulfonivorans*, proteins of interest intensity data part 1. Unidentified proteins are marked as N/A.*

Enzyme			LFQ Intensity								
Protein ID	Protein Name	Pathway	Control			Sulfur			Carbon/Sulfur		
			1	2	3	1	2	3	1	2	3
C6Y62_00305	Pdo	S	22.33	22.64	22.64	22.78	22.94	23.64	23.04	23.59	23.54
C6Y62_00310	Rhd-SqrB	S	27.91	27.82	27.77	27.55	27.44	27.41	27.38	27.41	27.45
C6Y62_00370	MxaI	MeOH	24.66	24.89	25.38	25.60	24.76	24.96	23.88	23.24	21.49
C6Y62_00375	MxaG	MeOH	29.47	29.06	29.51	29.29	29.33	29.29	26.11	25.89	27.20
C6Y62_00380	MxaJ	MeOH	31.40	31.26	31.05	31.16	31.13	30.97	28.47	27.83	29.69
C6Y62_00385	MxaF	MeOH	35.72	35.73	35.80	35.46	35.57	35.59	33.59	33.54	34.47
C6Y62_00835	MsuD	MSC	19.43	20.35	19.31	31.84	32.17	32.05	28.86	28.99	29.61
C6Y62_00880	SfnF	MSC	21.94	21.75	21.91	29.50	29.80	29.73	26.87	26.96	26.78
C6Y62_00885	SfnG	MSC	24.79	25.72	25.93	33.48	33.63	33.69	30.45	30.54	30.06
C6Y62_02365	FdhA	HCHO	28.11	28.03	27.67	27.82	27.77	27.95	28.59	28.66	28.39
C6Y62_02980	Fdh	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_04625	Acs	S	29.95	29.77	29.91	29.69	29.78	29.65	30.58	30.08	30.53
C6Y62_04710	FolD	HCHO	32.84	32.97	33.04	32.99	32.93	32.93	32.65	32.39	32.30
C6Y62_04850	Fhs	HCHO	33.24	33.23	33.10	33.06	33.28	33.09	32.87	32.34	33.06
C6Y62_05070	MetZ	S	29.38	29.43	29.25	29.21	29.36	29.24	28.57	28.24	28.79
C6Y62_05220	Fae	HCHO	30.04	30.30	30.52	30.34	30.18	30.34	29.42	29.82	28.65
C6Y62_05595	Fdh	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_08560	CysK	S	25.93	26.09	25.87	26.05	26.47	26.13	26.91	27.36	26.96
C6Y62_09260	catalase	MSC	28.27	28.66	28.57	29.26	28.96	29.07	31.01	30.58	30.40
C6Y62_09280	CysE	S	28.08	27.83	27.94	27.99	27.88	28.12	28.11	28.36	28.09
C6Y62_10250	FdhA	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_10660	SoxY	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_10665	SoxZ	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Supplementary Table S4.3-2ad: *Comparative proteomics of* H. sulfonivorans*, Experiment 2, proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			Comparative Analysis								
Protein ID	Protein Name	Pathway	Sulfur versus Control			Carbon/Sulfur versus Control			Carbon/Sulfur versus Sulfur		
			p-value	Welch's q-value	Fold Change	p-value	Welch's q-value	Fold Change	p-value	Welch's q-value	Fold Change
C6Y62_00305	Pdo	S	0.107	0.339	1.50	0.014	0.031	1.81	0.438	0.592	1.21
C6Y62_00310	Rhd-SqrB	S	0.003	0.071	0.77	0.001	0.005	0.75	0.329	0.475	0.96
C6Y62_00370	MxaI	MeOH	0.721	0.876	1.09	0.047	0.119	0.23	0.042	0.089	0.21
C6Y62_00375	MxaG	MeOH	0.790	0.912	0.97	0.002	0.018	0.13	0.002	0.029	0.13
C6Y62_00380	MxaJ	MeOH	0.267	0.497	0.90	0.010	0.052	0.17	0.011	0.054	0.19
C6Y62_00385	MxaF	MeOH	0.012	0.117	0.87	0.003	0.035	0.27	0.005	0.041	0.31
C6Y62_00835	MsuD	MSC	0.000	0.014	5128.90	0.000	0.000	703.15	0.000	0.009	0.14
C6Y62_00880	SfnF	MSC	0.000	0.000	224.21	0.000	0.000	32.07	0.000	0.000	0.14
C6Y62_00885	SfnG	MSC	0.000	0.041	277.77	0.000	0.002	29.20	0.000	0.004	0.11
C6Y62_02365	FdhA	HCHO	0.542	0.783	0.94	0.018	0.036	1.52	0.002	0.010	1.63
C6Y62_02980	Fdh	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_04625	Acs	S	0.061	0.231	0.89	0.036	0.098	1.44	0.013	0.053	1.62
C6Y62_04710	FolD	HCHO	0.991	0.997	1.00	0.014	0.034	0.70	0.009	0.050	0.70
C6Y62_04850	Fhs	HCHO	0.596	0.810	0.97	0.116	0.235	0.74	0.158	0.264	0.76
C6Y62_05070	MetZ	S	0.323	0.558	0.94	0.008	0.036	0.57	0.011	0.049	0.60
C6Y62_05220	Fae	HCHO	0.996	0.999	1.00	0.056	0.118	0.50	0.046	0.120	0.50
C6Y62_05595	Fdh	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_08560	CysK	S	0.157	0.380	1.19	0.002	0.015	2.16	0.011	0.021	1.81
C6Y62_09260	catalase	MSC	0.016	0.119	1.51	0.001	0.002	4.48	0.002	0.012	2.96
C6Y62_09280	CysE	S	0.664	0.855	1.03	0.105	0.145	1.18	0.163	0.203	1.14
C6Y62_10250	FdhA	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_10660	SoxY	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_10665	SoxZ	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Supplementary Table S4.3-1ba: *Comparative proteomics of* H. sulfonivorans*, Experiment 2, proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			Unique Peptides									MS/MS count								
Protein ID	Protein Name	Pathway	Control			Sulfur			Carbon/Sulfur			Control			Sulfur			Carbon/Sulfur		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
C6Y62_12280	SoxYZ	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_12500	FchC	HCHO	21	22	21	21	21	19	20	21	19	37	38	46	39	44	34	45	32	37
C6Y62_12505	FchD	HCHO	15	17	17	17	17	16	16	15	16	28	28	41	42	38	41	39	32	31
C6Y62_12510	FchA	HCHO	45	42	46	45	45	45	44	39	44	108	107	123	109	113	112	129	63	112
C6Y62_12515	FchB	HCHO	28	34	33	33	34	33	33	29	32	77	72	71	68	74	68	76	43	78
C6Y62_12520	XoxF	MeOH	5	6	2	3	4	3	7	6	6	4	2	1	2	1	0	8	0	6
C6Y62_12595	Fae	HCHO	14	15	15	15	16	16	13	14	14	75	108	99	116	118	118	119	94	110
C6Y62_12610	Mch	HCHO	18	17	18	19	20	21	18	18	16	42	37	39	41	47	44	37	30	44
C6Y62_12620	MtdB	HCHO	18	19	19	20	19	20	19	20	18	81	92	110	104	95	111	98	57	86
C6Y62_12675	CysJ	S	30	33	35	33	33	34	27	24	29	37	37	36	44	40	41	28	12	29
C6Y62_12695	CysNC	S	28	29	28	28	29	29	27	25	25	31	30	43	41	42	38	32	13	29
C6Y62_12700	CysD	S	14	13	14	15	16	13	12	12	13	23	17	19	18	28	24	19	6	17
C6Y62_12705	CysH	S	11	12	12	13	15	13	11	9	9	8	8	11	14	14	17	5	2	6
C6Y62_13025	Fae	HCHO	4	5	5	6	6	6	5	4	6	3	7	9	6	7	9	9	3	10
C6Y62_13185	SfnF	MSC	1	0	1	6	5	4	12	10	11	0	0	0	2	0	0	30	13	22
C6Y62_13190	SfnG	MSC	7	13	13	27	30	27	29	30	31	2	5	3	22	25	20	78	53	77
C6Y62_13200	DmoB	MSC	4	4	3	10	9	9	15	15	15	1	0	1	8	10	7	36	26	26
C6Y62_13210	DmoA	MSC	26	28	24	41	41	41	54	49	55	17	14	16	82	86	85	585	420	532
C6Y62_13220	SoxC	S	4	3	2	27	29	26	37	33	36	0	0	0	21	23	18	92	52	95
C6Y62_13225	SoxD	S	0	0	0	2	2	2	4	3	4	0	0	0	0	0	0	6	2	5
C6Y62_14215	SoxYZ	S	8	7	6	8	9	9	13	12	14	3	5	4	3	1	2	21	14	23
C6Y62_15095	CysC	S	3	3	5	3	4	4	4	2	6	2	1	2	0	1	2	2	0	3
C6Y62_15390	catalase	MSC	1	1	1	2	2	1	5	3	4	0	0	0	0	1	0	5	0	2
C6Y62_15815	CysK	S	13	16	14	16	18	18	16	15	16	16	16	21	27	44	31	31	12	29

Supplementary Table S4.3-2bb: *Comparative proteomics of* H. sulfonivorans*, Experiment 2, proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			Intensity								
Protein ID	Protein Name	Pathway	Control			Sulfur			Carbon/Sulfur		
			1	2	3	1	2	3	1	2	3
C6Y62_12280	SoxYZ	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_12500	FchC	HCHO	1.40E+09	1.39E+09	1.93E+09	1.03E+09	1.09E+09	9.88E+08	1.64E+09	6.60E+07	2.26E+09
C6Y62_12505	FchD	HCHO	1.30E+09	1.33E+09	1.71E+09	1.79E+09	1.34E+09	1.54E+09	1.66E+09	1.37E+08	1.94E+09
C6Y62_12510	FchA	HCHO	4.64E+09	3.94E+09	5.03E+09	3.32E+09	3.82E+09	3.60E+09	5.33E+09	9.66E+07	7.90E+09
C6Y62_12515	FchB	HCHO	4.19E+09	3.79E+09	4.18E+09	2.95E+09	3.19E+09	2.72E+09	4.92E+09	8.90E+07	7.51E+09
C6Y62_12520	XoxF	MeOH	1.65E+07	1.53E+07	8.72E+06	7.87E+06	1.10E+07	8.19E+06	3.84E+07	4.81E+05	3.58E+07
C6Y62_12595	Fae	HCHO	1.39E+10	1.57E+10	2.08E+10	1.41E+10	2.14E+10	1.76E+10	2.38E+10	8.78E+08	1.96E+10
C6Y62_12610	Mch	HCHO	4.59E+09	3.97E+09	4.52E+09	3.38E+09	3.65E+09	3.34E+09	3.12E+09	1.11E+08	4.00E+09
C6Y62_12620	MtdB	HCHO	5.73E+09	7.59E+09	9.58E+09	7.04E+09	7.06E+09	7.48E+09	7.86E+09	6.20E+08	8.52E+09
C6Y62_12675	CysJ	S	1.36E+09	9.67E+08	1.04E+09	1.10E+09	1.22E+09	1.08E+09	4.99E+08	1.78E+07	7.37E+08
C6Y62_12695	CysNC	S	8.08E+08	8.60E+08	9.34E+08	8.62E+08	1.09E+09	9.44E+08	5.17E+08	9.05E+06	9.24E+08
C6Y62_12700	CysD	S	4.68E+08	4.80E+08	7.77E+08	5.81E+08	7.08E+08	6.02E+08	2.48E+08	4.52E+06	4.06E+08
C6Y62_12705	CysH	S	1.06E+08	9.98E+07	1.38E+08	1.25E+08	1.77E+08	1.52E+08	6.40E+07	1.46E+06	8.80E+07
C6Y62_13025	Fae	HCHO	5.20E+07	6.93E+07	7.18E+07	4.76E+07	4.98E+07	4.70E+07	8.62E+07	3.56E+06	1.01E+08
C6Y62_13185	SfnF	MSC	1.10E+06	0.00	1.86E+05	1.20E+07	9.70E+06	1.38E+07	1.51E+09	2.97E+07	1.19E+09
C6Y62_13190	SfnG	MSC	3.46E+07	3.64E+07	4.84E+07	3.76E+08	3.55E+08	3.08E+08	5.18E+09	1.11E+08	7.29E+09
C6Y62_13200	DmoB	MSC	1.44E+07	1.31E+07	7.48E+06	2.02E+08	1.56E+08	1.42E+08	2.60E+09	7.63E+07	3.74E+09
C6Y62_13210	DmoA	MSC	3.62E+08	3.10E+08	2.31E+08	4.42E+09	4.76E+09	3.97E+09	6.51E+10	3.25E+09	8.52E+10
C6Y62_13220	SoxC	S	1.17E+07	6.19E+06	7.74E+06	2.77E+08	2.57E+08	2.33E+08	3.75E+09	5.98E+07	5.79E+09
C6Y62_13225	SoxD	S	0.00	0.00	0.00	5.41E+06	5.69E+06	5.35E+06	1.77E+08	3.45E+06	1.83E+08
C6Y62_14215	SoxYZ	S	3.22E+07	1.86E+07	2.06E+07	3.12E+07	3.51E+07	2.97E+07	2.72E+08	4.26E+06	4.20E+08
C6Y62_15095	CysC	S	1.07E+07	9.58E+06	1.35E+07	4.60E+06	5.56E+06	7.25E+06	1.92E+07	1.59E+05	3.94E+07
C6Y62_15390	catalase	MSC	5.26E+06	4.03E+06	5.31E+06	6.86E+06	6.97E+06	2.79E+06	5.19E+07	2.18E+05	2.21E+07
C6Y62_15815	CysK	S	3.75E+08	2.87E+08	5.02E+08	7.95E+08	1.00E+09	8.59E+08	8.74E+08	2.51E+07	9.81E+08

Supplementary Table S4.3-2bc: *Comparative proteomics of* H. sulfonivorans*, Experiment 2, proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			LFQ Intensity								
Protein ID	Protein Name	Pathway	Control			Sulfur			Carbon/Sulfur		
			1	2	3	1	2	3	1	2	3
C6Y62_12280	SoxYZ	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C6Y62_12500	FchC	HCHO	30.55	30.46	30.52	30.29	30.23	30.20	30.61	31.16	30.48
C6Y62_12505	FchD	HCHO	31.06	30.97	30.84	30.97	31.00	30.98	30.84	31.52	30.48
C6Y62_12510	FchA	HCHO	32.07	32.04	32.03	31.85	32.06	32.06	32.23	31.55	32.34
C6Y62_12515	FchB	HCHO	31.91	31.92	31.84	31.77	31.86	31.73	32.13	31.74	32.32
C6Y62_12520	XoxF	MeOH	24.09	23.80	23.66	23.59	23.47	23.53	25.16	24.38	24.79
C6Y62_12595	Fae	HCHO	33.81	34.32	34.10	34.48	34.59	34.59	34.44	34.53	33.45
C6Y62_12610	Mch	HCHO	32.12	31.99	31.81	31.92	31.92	31.88	31.51	31.79	31.29
C6Y62_12620	MtdB	HCHO	33.02	33.25	33.30	33.38	33.24	33.37	33.32	33.65	32.55
C6Y62_12675	CysJ	S	29.98	30.14	30.08	30.44	30.41	30.40	29.22	29.24	29.06
C6Y62_12695	CysNC	S	29.60	29.75	29.70	29.85	30.16	30.02	29.18	29.02	29.57
C6Y62_12700	CysD	S	28.76	29.00	29.32	29.34	29.63	29.44	28.12	28.09	28.18
C6Y62_12705	CysH	S	27.05	26.89	27.07	27.27	27.40	27.36	25.98	25.99	25.98
C6Y62_13025	Fae	HCHO	25.88	26.14	26.18	25.88	25.81	25.65	26.37	27.37	26.10
C6Y62_13185	SfnF	MSC	22.56	20.54	21.85	22.98	22.58	23.17	30.51	30.42	30.05
C6Y62_13190	SfnG	MSC	25.42	25.68	25.65	28.50	28.39	28.31	32.41	32.30	32.42
C6Y62_13200	DmoB	MSC	23.22	23.12	22.69	27.61	27.21	27.12	31.58	31.56	31.48
C6Y62_13210	DmoA	MSC	29.36	29.21	28.79	33.01	33.06	32.87	36.55	36.22	36.26
C6Y62_13220	SoxC	S	21.89	21.88	21.91	27.46	27.24	27.20	32.12	30.28	32.26
C6Y62_13225	SoxD	S	21.01	23.67	22.23	23.19	23.16	23.09	26.06	27.55	27.69
C6Y62_14215	SoxYZ	S	25.15	24.88	24.68	25.35	25.51	25.39	28.11	27.83	28.02
C6Y62_15095	CysC	S	23.49	23.86	23.15	21.62	23.19	23.26	24.39	23.74	24.00
C6Y62_15390	catalase	MSC	20.40	22.26	22.64	21.24	24.05	23.78	25.20	23.98	24.37
C6Y62_15815	CysK	S	28.76	28.61	28.74	30.08	30.06	29.98	29.79	29.73	29.45

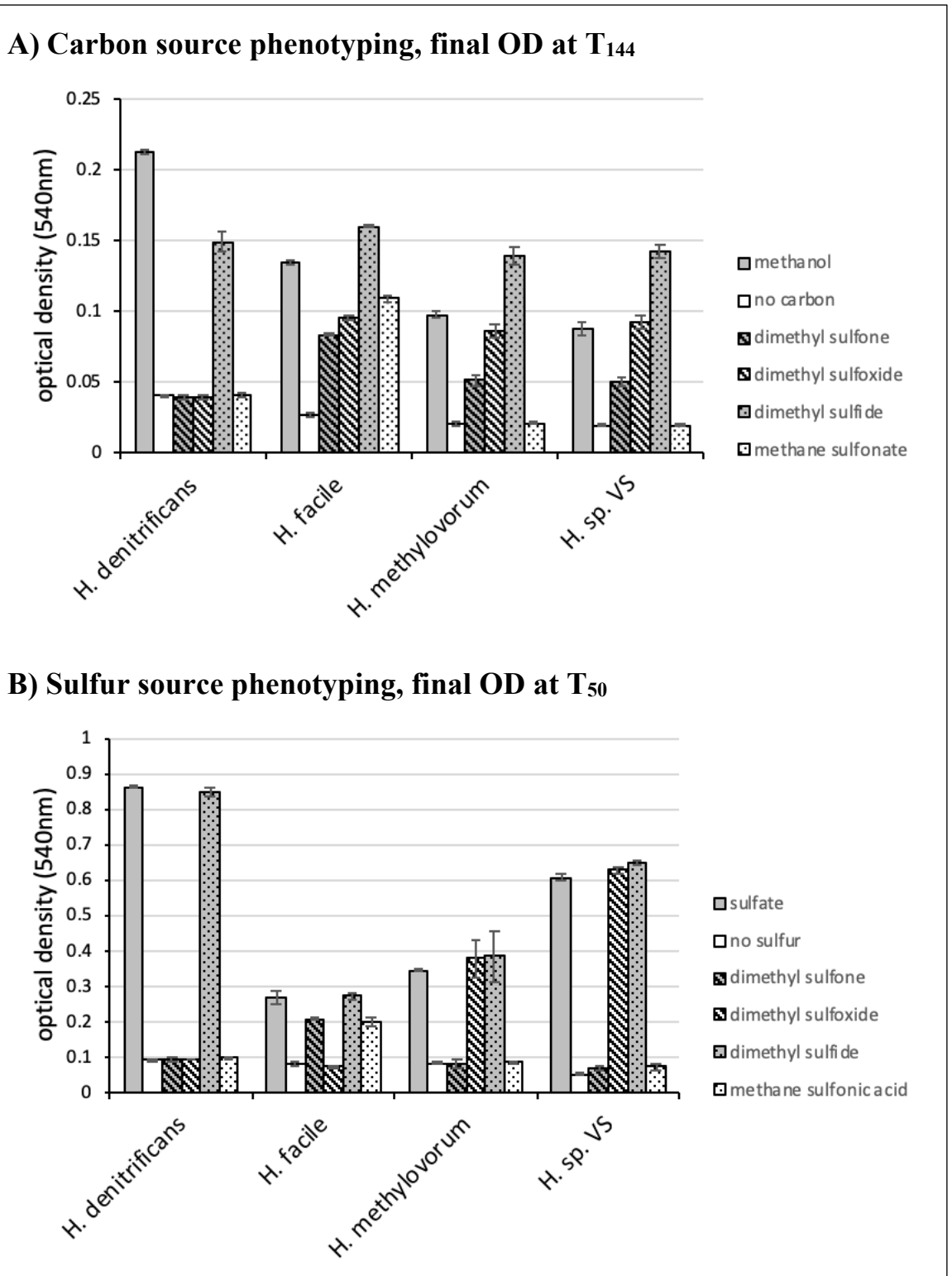


FIGURE S5.2-1: Hyphomicrobium *MSC utilisation phenotyping, additional figures.* (A) OD_{540} of *Hyphomicrobium* species cultures on various carbon sources at final sampling after 144 h fed-batch cultivation, supplemented with a sulfur source of Na_2SO_4. (B) OD_{540} of *Hyphomicrobium* species cultures on various sulfur sources at final sampling after 50 h batch cultivation, supplemented with a carbon source of MeOH.

```
Draft  1                                                CAACTTGAGAGTTTGATCCTGGCTCAG  27

Draft  28    AACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAACGCCCCGCAAGGGGAGTGGCA  87
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1     AACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAACGCCCCGCAAGGGGAGTGGCA  60

Draft  88    GACGGGTGAGTAACACGTGGGAACCTTCCCTATAGTACGGAATAGCCCAGGGAAACTTGG  147
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    GACGGGTGAGTAACACGTGGGAACCTTCCCTATAGTACGGAATAGCCCAGGGAAACTTGG  120

Draft  148   AGTAATACCGTATACGCCCGAGAGGGGAAAGAATTTCGCTATAGGATGGGCCCGCGTAGG  207
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121   AGTAATACCGTATACGCCCGAGAGGGGAAAGAATTTCGCTATAGGATGGGCCCGCGTAGG  180

Draft  208   ATTAGCTAGTTGGTGAGGTAATGGCTCACCAAGGCGACGATCCTTAGCTGGTTTGAGAGA  267
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181   ATTAGCTAGTTGGTGAGGTAATGGCTCACCAAGGCGACGATCCTTAGCTGGTTTGAGAGA  240

Draft  268   ACGACCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGG  327
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241   ACGACCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGG  300

Draft  328   AATATTGGACAATGGGCGAAAGCCTGATCCAGCCATGCCGCGTGAGTGATGAAGGCCTTA  387
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301   AATATTGGACAATGGGCGAAAGCCTGATCCAGCCATGCCGCGTGAGTGATGAAGGCCTTA  360

Draft  388   GGGTTGTAAAGCTCTTTTGCCGGGGACGATAATGACGGTACCCGGAGAATAAGTCCCGGC  447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361   GGGTTGTAAAGCTCTTTTGCCGGGGACGATAATGACGGTACCCGGAGAATAAGTCCCGGC  420

Draft  448   TAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGGACTAGCGTTGTTCGGAATCACTGG  507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421   TAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGGACTAGCGTTGTTCGGAATCACTGG  480

Draft  508   GCGTAAAGCGCACGTAGGCGGATTTATAAGTCAGGGGTGAAATCCCGGGGCTCAACCTCG  567
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  481   GCGTAAAGCGCACGTAGGCGGATTTATAAGTCAGGGGTGAAATCCCGGGGCTCAACCTCG  540

Draft  568   GAACTGCCTTTGATACTGTGAATCTTGAGTCCGATAGAGGTGGGTGGAATTCCTAGTGTA  627
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  541   GAACTGCCTTTGATACTGTGAATCTTGAGTCCGATAGAGGTGGGTGGAATTCCTAGTGTA  600

Draft  628   GAGGTGAAATTCGTAGATATTAGGAAGAACACCGGTGGCGAAGGCGGCCCACTGGATCGG  687
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  601   GAGGTGAAATTCGTAGATATTAGGAAGAACACCGGTGGCGAAGGCGGCCCACTGGATCGG  660

Draft  688   TACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCA  747
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  661   TACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCA  720
```

Supplementary Figure S5.3-1a: *BLASTn results of the 16S RNA gene sequence from* H. methylovorum *Bras1 draft 2018 (Draft) versus the* H. methylovorum *NBRC 14180* strain partial *16S RNA gene (Sbject, NCBI Accession NR_113655.1)).*

```
Draft  748   CGCCGTAAACGATGGATGCTAGCCGTCGGATAGCTTGCTATTCGGTGGCGCAGCTAACGC  807
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  721   CGCCGTAAACGATGGATGCTAGCCGTCGGATAGCTTGCTATTCGGTGGCGCAGCTAACGC  780

Draft  808   ATTAAGCATCCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGG  867
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  781   ATTAAGCATCCCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAAGGAATTGACGGGG  840

Draft  868   GCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGACGCAACGCGAAGAACCTTACCAGC  927
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  841   GCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGACGCAACGCGAAGAACCTTACCAGC  900

Draft  928   TCTTGACATTCACTGATCGCCTGGAGAGATCCGGGAATTCCAGCAATGGACAGTGGGACA  987
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  901   TCTTGACATTCACTGATCGCCTGGAGAGATCCGGGAATTCCAGCAATGGACAGTGGGACA  960

Draft  988   GGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAG  1047
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  961   GGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAG  1020

Draft  1048  CGCAACCCTCGCCATTAGTTGCCATCATTTAGTTGGGCACTCTAGTGGGACTGCCGGTGA  1107
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1021  CGCAACCCTCGCCATTAGTTGCCATCATTTAGTTGGGCACTCTAGTGGGACTGCCGGTGA  1080

Draft  1108  TAAGCCGGAGGAAGGTGGGGATGACGTCAAGTCATCATGGCCCTTACGGGCTGGGCTACA  1167
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1081  TAAGCCGGAGGAAGGTGGGGATGACGTCAAGTCATCATGGCCCTTACGGGCTGGGCTACA  1140

Draft  1168  CACGTGCTACAATGGCGGTGACAATGCGCAGCCACCTAGTAATAGGGAGCTAATCGCAAA  1227
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1141  CACGTGCTACAATGGCGGTGACAATGCGCAGCCACCTAGTAATAGGGAGCTAATCGCAAA  1200

Draft  1228  AAGCCGTCTCAGTTCAGATTGAGGTCTGCAACTCGACCTCATGAAGTCGGAATCGCTAGT  1287
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1201  AAGCCGTCTCAGTTCAGATTGAGGTCTGCAACTCGACCTCATGAAGTCGGAATCGCTAGT  1260

Draft  1288  AATCGCGCATCAGCATGGCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCA  1347
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1261  AATCGCGCATCAGCATGGCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCA  1320

Draft  1348  CACCATGGGAGTTGGTCTTACCCTAAAACGGTGCGCTAACCGCAAGGAGGCAGCCGGCCA  1407
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1321  CACCATGGGAGTTGGTCTTACCCTAAAACGGTGCGCTAACCGCAAGGAGGCAGCCGGCCA  1380

Draft  1408  CGGTAAGGTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAACCTGCGGCT  1457
             ||||||||||||||||||||||||||
Sbjct  1381  CGGTAAGGTCAGCGACTGGGGTGAAG                                    1406

Draft  1458  GGATCACCTCCTTTCTA                                              1474
```

Supplementary Figure S5.3-1b: *BLASTn results of the 16S RNA gene sequence from* H. methylovorum *Bras1 draft 2018 (Draft) versus the* H. methylovorum *NBRC 14180* strain's partial *16S RNA gene (Sbjct, NCBI Accession NR_113655.1).*

Supplementary Table S5.4-1: *BLAST search of the* H. denitrificans *strain ATCC 51888 for genes of methanol metabolism.*

METHANOL METABOLISM									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
MXO system MeOH dehydrogenase	subunit 1	MxaF	UNIPROT/P16027	*Methylobacterium extorquens*	**HDEN_RS06525**	**PQQ-dependent dehydrogenase, methanol/ethanol family**	**80%**	**1015**	**0.0**
					HDEN_RS06435	PQQ-dependent dehydrogenase, methanol/ethanol family	48%	545	0.0
					HDEN_RS09180	PQQ-dependent dehydrogenase, methanol/ethanol	48%	540	0.0
					HDEN_RS14080	PQQ-dependent dehydrogenase, methanol/ethanol family	47%	535	0.0
					HDEN_RS06435	PQQ-dependent dehydrogenase, methanol/ethanol family	47%	518	3e-177
					HDEN_RS07935	PQQ-dependent dehydrogenase, methanol/ethanol family	35%	251	3e-75
					HDEN_RS06525	PQQ-dependent dehydrogenase, methanol/ethanol family	31%	241	3e-71
	subunit 2	MxaI	UNIPROT/P14775	*Methylobacterium extorquens*	**HDEN_RS06510**	**methanol dehydrogenase**	**82%**	**125**	**7e-39**
MXO system protein		MxaJ	UNIPROT/P16028	*Methylobacterium extorquens*	**HDEN_RS06520**	**methanol oxidation system protein MoxJ**	**52%**	**249**	**2e-81**
					HDEN_RS07945	quinoprotein dehydrogenase-associated putative ABC transporter substrate-binding protein	30%	93.2	9e-23
MXO system cytochrome c-L		MxaG	UNIPROT/P14774	*Methylobacterium extorquens*	**HDEN_RS06515**	**cytochrome c(L) MoxG**	**61%**	**202**	**3e-06**
					HDEN_RS03890	cytochrome c, class I	28%	42.4	4e-17
					HDEN_RS10100	cytochrome c	40%	38.9	9e-05
(continued)									

Supplementary Table S5.4-1b: *BLAST search of the* H. methylovorum *strain Bras1 2018 draft for genes of methanol metabolism (continued).*

METHANOL METABOLISM *(continued)*								
REFERENCE SEQUENCE				BLAST RESULTS				
Enzyme	Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
MeOH dehydrogenase	XoxF	UNIPROT/C5ATJ3	*Methylobacterium extorquens*	**HDEN_RS09180**	**PQQ-dependent dehydrogenase, methanol/ethanol family**	**75%**	**842**	**0.0**
				HDEN_RS07935	**PQQ-dependent dehydrogenase, methanol/ethanol family**	**72%**	**823**	**0.0**
				HDEN_RS06435	**PQQ-dependent dehydrogenase, methanol/ethanol family**	**73%**	**822**	**0.0**
				HDEN_RS14080	PQQ-dependent dehydrogenase, methanol/ethanol family	67%	765	0.0
				HDEN_RS06525	PQQ-dependent dehydrogenase, methanol/ethanol family	51%	529	0.0
				HDEN_RS01645	PQQ-dependent dehydrogenase, methanol/ethanol family	39%	318	1e-100
				HDEN_RS03895	PQQ-dependent dehydrogenase, methanol/ethanol family	33%	251	2e-75
quinoprotein alcohol dehydrogenase	ExaA	UNIPROT/Q9Z4J7	*Pseudomonas aeruginosa*	HDEN_RS01645	PQQ-dependent dehydrogenase, methanol/ethanol family	55%	605	0.0
				HDEN_RS03895	PQQ-dependent dehydrogenase, methanol/ethanol	34%	304	5e-95
				HDEN_RS09180	PQQ-dependent dehydrogenase, methanol/ethanol family	36%	300	4e-93
				HDEN_RS14080	PQQ-dependent dehydrogenase, methanol/ethanol family	37%	297	3e-92
				HDEN_RS06435	PQQ-dependent dehydrogenase, methanol/ethanol family	36%	297	5e-92
				HDEN_RS07935	PQQ-dependent dehydrogenase, methanol/ethanol family	36%	293	2e-90
				HDEN_RS06525	PQQ-dependent dehydrogenase, methanol/ethanol family	34%	266	2e-80

Supplementary Table S5.4-2a: *BLAST search of the* H. methylovorum *strain Bras1 2018 draft for genes of methanol metabolism.*

METHANOL METABOLISM									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
MXO system MeOH dehydrogenase	subunit 1	MxaF	UNIPROT/P16027	*Methylobacterium extorquens*	**DLM45_03560**	**PQQ-dependent dehydrogenase, methanol/ethanol family**	**80%**	**1016**	**0.0**
					DLM45_00110	PQQ-dependent dehydrogenase, methanol/ethanol family	48%	538	0.0
					DLM45_14260	PQQ-dependent dehydrogenase, methanol/ethanol	47%	516	1e-176
					DLM45_03670	PQQ-dependent dehydrogenase, methanol/ethanol family	41%	401	3e-132
					DLM45_05150	PQQ-dependent dehydrogenase, methanol/ethanol family	39%	376	5e-122
	subunit 2	MxaI	UNIPROT/P14775	*Methylobacterium extorquens*	**DLM45_03575**	**methanol dehydrogenase**	**75%**	**116**	**4e-36**
MXO system protein		MxaJ	UNIPROT/P16028	*Methylobacterium extorquens*	**DLM45_03565**	**methanol oxidation system protein MoxJ**	**52%**	**253**	**9e-84**
					DLM45_05145	quinoprotein dehydrogenase-associated putative ABC transporter substrate-binding protein	32%	105	3e-26
					DLM45_03665	quinoprotein dehydrogenase-associated putative ABC transporter substrate-binding protein	31%	102	7e-26
					DLM45_00100	quinoprotein dehydrogenase-associated putative ABC transporter substrate-binding protein	27%	81.6	1e-19
MXO system cytochrome c-L		MxaG	UNIPROT/P14774	*Methylobacterium extorquens*	**DLM45_03570**	**cytochrome c(L) MoxG**	**61%**	**187**	**4e-61**
					DLM45_14535	cytochrome c, class I	37%	72.8	4e-17
					DLM45_03660	cytochrome c	31%	42.7	3e-07
(continued)									

Table S5.4-2b: *BLAST search of the* H. methylovorum *strain Bras1 2018 draft for genes of methanol metabolism (continued).*

METHANOL METABOLISM *(continued)*								
REFERENCE SEQUENCE				BLAST RESULTS				
Enzyme	Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
MeOH dehydrogenase	XoxF	UNIPROT/C5ATJ3	*Methylobacterium extorquens*	**DLM45_14260**	**PQQ-dependent dehydrogenase, methanol/ethanol family**	**74%**	**820**	**0.0**
				DLM45_00110	**PQQ-dependent dehydrogenase, methanol/ethanol family**	**72%**	**816**	**0.0**
				DLM45_03560	PQQ-dependent dehydrogenase, methanol/ethanol family	50%	531	0.0
				DLM45_03670	PQQ-dependent dehydrogenase, methanol/ethanol family	43%	426	3e-142
quinoprotein alcohol dehydrogenase	ExaA	UNIPROT/Q9Z4J7	*Pseudomonas aeruginosa*	DLM45_00110	PQQ-dependent dehydrogenase, methanol/ethanol family	36%	289	7e-90
				DLM45_14260	PQQ-dependent dehydrogenase, methanol/ethanol	35%	287	2e-88
				DLM45_03560	PQQ-dependent dehydrogenase, methanol/ethanol family	35%	271	6e-83
				DLM45_03670	PQQ-dependent dehydrogenase, methanol/ethanol family	34%	270	3e-82
				DLM45_03865	pyrrolo-quinoline quinone	26%	55.5	1e-09

Supplementary Table S5.4-3a: *BLAST search of the* H. denitrificans *strain Bras1 2018 draft for genes of methylated sulfur compound metabolism.*

METHYLATED SULFUR COMPOUND METABOLISM									
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Interconversion of dimethyl sulfur compounds									
DMSO reductase	subunit A	DmsA	UNIPROT/P18775	*Escherichia coli*	HDEN_RS07070	molybdopterin oxidoreductase	28%	223	9e-63
					HDEN_RS12165	molybdopterin oxidoreductase	21%	92.8	4e-20
	subunit B	DmsB	UNIPROT/P18776	*Escherichia coli*	HDEN_RS04565	nitrate reductase subunit beta	45%	92.4	1e-22
	subunit C	DmsC	UNIPROT/P18777	*Escherichia coli*	*no significant hits*				
DMS dehydrogenase	subunit A	DdhA	UNIPROT/Q8GPG4	*Rhodovulum sulfidophilum*	HDEN_RS04560	nitrate reductase subunit alpha	32%	185	7e-49
					HDEN_RS07070	molybdopterin oxidoreductase family protein	29%	78.2	1e-15
	subunit B	DdhB	UNIPROT/Q8GPG3	*Rhodovulum sulfidophilum*	HDEN_RS04565	nitrate reductase subunit beta	36%	247	9e-78
	subunit C	DdhC	UNIPROT/Q8GPG1	*Rhodovulum sulfidophilum*	*no significant hits*				
Dimethylsulfide oxidation pathway									
DMS monooxygenase large subunit		DmoA	NCBI/ADU77278.1	*Hyphomicrobium sulfonivorans*	*no significant hits*				
MT oxidase		MtoX	NCBI/ATJ26742.1	*Hyphomicrobium* sp. VS	**HDEN_RS03620**	**selenium-binding protein**	**77%**	**666**	**0.0**
Dimethylsulfone oxidation pathway									
MSA monooxygenase	subunit A	MsmA	UNIPROT/Q9X404	*Methylosulfonomonas methylovora*	*no significant hits*				
	subunit B	MsmB	UNIPROT/Q9X405	*Methylosulfonomonas methylovora*	*no significant hits*				
	subunit C	MsmC	UNIPROT/Q9X405	*Methylosulfonomonas methylovora*	HDEN_RS14590	nitrite reductase (NAD(P)H) small subunit	28%	44.3	7e-08
	subunit D	MsmD	UNIPROT/Q9X406	*Methylosulfonomonas methylovora*	HDEN_RS03575	oxidoreductase FAD/NAD(P)-binding domain-containing protein	26%	94.4	1e-22
(continued)									

Supplementary Table S5.4-3b: *BLAST search of the* H. denitrificans *strain ATCC 51888 genome for genes of methylated sulfur compound metabolism (continued).*

METHYLATED SULFUR COMPOUND METABOLISM *(continued)*									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Dimethylsulfone oxidation pathway (continued)									
$DMSO_2$ monooxygenase	large subunit	SfnG	UNIPROT/Q3KC85	*Pseudomonas fluorescens*	HDEN_RS08615	LLM class flavin-dependent oxidoreductase	23%	41.2	6e-05
	small subunit	SfnF	UNIPROT/Q3K9A2	*Pseudomonas fluorescens*	*no significant hits*				
MSA monooxygenase	large subunit	MsuD	UNIPROT/Q9I1C2	*Pseudomonas aeruginosa*	HDEN_RS08615	LLM class flavin-dependent oxidoreductase	30%	40.0	1e-04
	small subunit	MsuE	UNIPROT/Q88J85	*Pseudomonas putida*	*no significant hits*				
alkanesulfonate monooxygenase	large subunit	SsuD	UNIPROT/Q8FJ93	*Escherichia coli*	HDEN_RS08615	LLM class flavin-dependent oxidoreductase	25%	59.3	1e-10
	small subunit	SsuE	UNIPROT/P80644	*Escherichia coli*	*no significant hits*				
DMS production pathway									
MT S-methyltransferase		MddA	UNIPROT/ A0A0F6P9C0	*Pseudomonas deceptionensis*	HDEN_RS10980	Isoprenylcysteine carboxylmethyltransferase family protein	35%	32.0	0.022
MT:CoM methyltransferase	subunit A	MtsA	UNIPROT/Q48924	*Methanosarcina barkeri*	*no significant hits*				
	subunit B	MtsB	UNIPROT/Q48925	*Methanosarcina barkeri*	HDEN_RS08840	methionine synthase	29%	79.0	2e-17
L-methionine gamma-lyase		MdeA	UNIPROT/P13254	*Pseudomonas_putida*	HDEN_RS07495	O-acetylhomoserine aminocarboxypropyltransferase	41%	286	5e-93
					HDEN_RS00130	O-succinylhomoserine sulfhydrylase	39%	249	5e-79
					HDEN_RS05360	cystathionine beta-lyase	32%	154	6e-43

Supplementary Table S5.4-4a: *BLAST search of the* H. methylovorum *strain Bras1 2018 draft for genes of methylated sulfur compound metabolism.*

METHYLATED SULFUR COMPOUND METABOLISM									
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Interconversion of dimethyl sulfur compounds									
DMSO reductase	subunit A	DmsA	UNIPROT/P18775	*Escherichia coli*	DLM45_12250	molybdopterin oxidoreductase	29%	224	3e-63
					DLM45_15805	molybdopterin oxidoreductase	28%	219	2e-61
	subunit B	DmsB	UNIPROT/P18776	*Escherichia coli*	DLM45_01800	DeoR/GlpR transcriptional regulator	45%	30.4	0.012
	subunit C	DmsC	UNIPROT/P18777	*Escherichia coli*	*no significant hits*				
DMS dehydrogenase	subunit A	DdhA	UNIPROT/Q8GPG4	*Rhodovulum sulfidophilum*	DLM45_12250	molybdopterin oxidoreductase	26%	82.4	5e-17
					DLM45_15805	molybdopterin oxidoreductase	21%	77.4	2e-15
	subunit B	DdhB	UNIPROT/Q8GPG3	*Rhodovulum sulfidophilum*	*no significant hits*				
	subunit C	DdhC	UNIPROT/Q8GPG1	*Rhodovulum sulfidophilum*	*no significant hits*				
Dimethylsulfide oxidation pathway									
DMS monooxygenase large subunit		DmoA	NCBI/ADU77278.1	*Hyphomicrobium sulfonivorans*	**DLM45_02930**	**5,10-methylene H_4MPT reductase**	**81%**	**807**	**0.0**
					DLM45_02945	$DMSO_2$ monooxygenase SfnG	25%	63.9	1e-12
MT oxidase		MtoX	NCBI/ATJ26742.1	*Hyphomicrobium* sp. VS	**DLM45_12340**	**selenium-binding protein**	**99%**	**840**	**0.0**
Dimethylsulfone oxidation pathway									
MSA monooxygenase	subunit A	MsmA	UNIPROT/Q9X404	*Methylosulfonomonas methylovora*	**DLM45_05095**	**methanesulfonate monooxygenase**	**85%**	**757**	**0.0**
	subunit B	MsmB	UNIPROT/Q9X405	*Methylosulfonomonas methylovora*	**DLM45_05100**	**hypothetical protein**	**61%**	**226**	**3e-76**
	subunit C	MsmC	UNIPROT/Q9X405	*Methylosulfonomonas methylovora*	**DLM45_05105**	**(2Fe-2S)-binding protein**	**64%**	**159**	**1e-51**
	subunit D	MsmD	UNIPROT/Q9X406	*Methylosulfonomonas methylovora*	**DLM45_05110**	**oxidoreductase**	**56%**	**380**	**1e-130**
					DLM45_12320	oxidoreductase	26%	94.7	8e-23
(continued)									

Supplementary Table S5.4-4b: *BLAST search of the* H. methylovorum *strain Bras1 2018 draft for genes of methylated sulfur compound metabolism (continued).*

METHYLATED SULFUR COMPOUND METABOLISM *(continued)*									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Dimethylsulfone oxidation pathway (continued)									
$DMSO_2$ monooxygenase	large subunit	SfnG	UNIPROT/Q3KC85	*Pseudomonas fluorescens*	**DLM45_02945**	**$DMSO_2$ monooxygenase SfnG**	**61%**	**452**	**1e-159**
					DLM45_02930	5,10-methylene t H_4MPT reductase	30%	81.6	1e-18
					DLM45_00805	LLM class flavin-dependent oxidoreductase	24%	51.6	5e-09
	small subunit	SfnF	UNIPROT/Q3K9A2	*Pseudomonas fluorescens*	**DLM45_02940**	**hypothetical protein**	**40%**	**114**	**1e-31**
MSA monooxygenase	large subunit	MsuD	UNIPROT/Q9I1C2	*Pseudomonas aeruginosa*	DLM45_02945	$DMSO_2$ monooxygenase SfnG	31%	117	5e-31
					DLM45_02930	5,10-methylene t H_4MPT reductase	31%	47.8	1e-07
					DLM45_00805	LLM class flavin-dependent oxidoreductase	27%	37.0	3e-04
	small subunit	MsuE	UNIPROT/Q88J85	*Pseudomonas putida*	DLM45_02940	hypothetical protein	38%	119	1e-33
alkanesulfonate monooxygenase	large subunit	SsuD	UNIPROT/Q8FJ93	*Escherichia coli*	DLM45_02945	dimethyl sulfone monooxygenase SfnG	31%	117	4e-31
					DLM45_00805	LLM class flavin-dependent oxidoreductase	26%	62.8	2e-12
					DLM45_02930	5,10-methylene H_4MPT reductase	24%	53.1	3e-09
	small subunit	SsuE	UNIPROT/P80644	*Escherichia coli*	DLM45_02940	hypothetical protein	31%	67.0	5e-15
DMS production pathway									
MT S-methyltransferase		MddA	UNIPROT/ A0A0F6P9C0	*Pseudomonas deceptionensis*	DLM45_03285	Isoprenylcysteine carboxylmethyltransferase family protein	35%	37.7	5e-05
MT:CoM methyltransferase	subunit A	MtsA	UNIPROT/Q48924	*Methanosarcina barkeri*	*no significant hits*				
	subunit B	MtsB	UNIPROT/Q48925	*Methanosarcina barkeri*	DLM45_01040	methionine synthase	28%	75.9	5e-17
L-methionine gamma-lyase		MdeA	UNIPROT/P13254	*Pseudomonas_putida*	DLM45_00235	O-acetylhomoserine aminocarboxypropyltransferase	41%	281	1e-91
					DLM45_09555	O-succinylhomoserine sulfhydrylase, MetZ	40%	273	5e-88

Supplementary Table S5.4-5a: *BLAST search of the* H. denitrificans *strain ATCC 51888 genome for genes of formaldehyde metabolism.*

FORMALDEHYDE METABOLISM									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Formaldehyde oxidation									
formaldehyde dehydrogenase		FdhA	UNIPROT/P46154	*Pseudomonas putida*	HDEN_RS00485	alcohol dehydrogenase	34%	100	2e-24
					HDEN_RS02725	NADPH:quinone oxidoreductase	34%	40.8	1e-04
Glutathione pathway									
GSH-dependent formaldehyde activating enzyme		Gfa	UNIPROT/Q51669	*Paracoccus denitrificans*	*no significant hits*				
hydroxymethyl-GSH dehydrogenase		FlhA	UNIPROT/P45382	*Paracoccus denitrificans*	HDEN_RS00485	alcohol dehydrogenase	25%	48.9	2e-07
					HDEN_RS02725	NADPH:quinone oxidoreductase	26%	45.1	3e-06
Tetrahydromethanopterin pathway									
formaldehyde-activating enzyme		Fae	UNIPROT/Q9FA38	*Methylobacterium extorquens*	**HDEN_RS07280**	**formaldehyde-activating enzyme**	**79%**	**278**	**1e-96**
					HDEN_RS09220	**formaldehyde-activating enzyme**	**36%**	**80.5**	**3e-21**
					HDEN_RS10485	**formaldehyde-activating enzyme**	**29%**	**78.2**	**5e-19**
methylene-H_4MPT dehydrogenase A		MtdA	UNIPROT/P55818	*Methylobacterium extorquens*	HDEN_RS07305	methylene-H_4MPT dehydrogenase	31%	65.5	3e-13
NADP-dependent methylene-H_4MPT dehydrogenase B		MtdB	UNIPROT/O85012	*Methylobacterium extorquens*	**HDEN_RS07305**	**methylene-H_4MPT dehydrogenase**	**52%**	**277**	**2e-92**
methenyl-H_4MPT cyclohydrolase		Mch	UNIPROT/O85014	*Methylobacterium extorquens*	**HDEN_RS07295**	**methenyl-H_4MPT cyclohydrolase**	**57%**	**374**	**9e-130**
formyltransferase/ hydrolase complex	subunit A	FhcA	UNIPROT/C5B137	*Methylobacterium extorquens*	**HDEN_RS07900**	**formyl-MFR dehydrogenase subunit A**	**43%**	**417**	**8e-140**
	subunit B	FhcB	UNIPROT/C5B138	*Methylobacterium extorquens*	**HDEN_RS07905**	**formyl-MFR dehydrogenase**	**27%**	**70.1**	**3e-14**
	subunit C	FchC	UNIPROT/C5B135	*Methylosulfonomonas methylovora*	**HDEN_RS07890**	**Formyl-MFR dehydrogenase subunit C**	**48%**	**178**	**5e-55**
	Subunit D	FchD	UNIPROT/Q49118	*Methylosulfonomonas methylovora*	**HDEN_RS07895**	**Formyl-MFR dehydrogenase subunit D**	**60%**	**359**	**3e-124**
(continued)									

Supplementary Table S5.4-5b: *BLAST search of the* H. denitrificans *strain ATCC 51888 genome for genes of formaldehyde metabolism (continued).*

FORMALDEHYDE METABOLISM *(continued)*								
REFERENCE SEQUENCE				BLAST RESULTS				
Enzyme	Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Formate oxidation								
formate dehydrogenase	Fdh	UNIPROT/P33160	*Pseudomonas* sp. 101	**HDEN_RS02970**	**NAD-dependent formate dehydrogenase**	**79%**	**660**	**0.0**
				HDEN_RS01380	D-glycerate dehydrogenase	30%	122	2e-32
				HDEN_RS17390	D-glycerate dehydrogenase	32%	115	1e-28
Tetrahydrofolate pathway								
formate-THF ligase	Fhs	UNIPROT/Q83WS0	*Methylobacterium extorquens*	**HDEN_RS15500**	**formate H_4F ligase**	**68%**	**709**	**0.0**
				HDEN_RS00520	**formate H_4F ligase**	**67%**	**704**	**0.0**
bifunctional methenyl-H_4F cyclohydrolase/ methylene-H_4F dehydrogenase	FolD	UNIPROT/P24186	*Escherichia coli*	**HDEN_RS15885**	**bifunctional methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase, FolD**	**56%**	**271**	**2e-90**
				HDEN_RS00515	**bifunctional methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase, FolD**	**49%**	**243**	**3e-79**

Supplementary Table S5.4-6a: *BLAST search of the* H. methylovorum *strain Bras1 2018 draft for genes of formaldehyde metabolism.*

FORMALDEHYDE METABOLISM									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Formaldehyde oxidation									
formaldehyde dehydrogenase		FdhA	UNIPROT/P46154	*Pseudomonas putida*	DLM45_11985	alcohol dehydrogenase	29%	35.0	0.006
Glutathione pathway									
GSH-dependent formaldehyde activating enzyme		Gfa	UNIPROT/Q51669	*Paracoccus denitrificans*	*no significant hits*				
hydroxymethyl-GSH dehydrogenase		FlhA	UNIPROT/P45382	*Paracoccus denitrificans*	DLM45_11985	alcohol dehydrogenase	33%	42.4	2e-05
Tetrahydromethanopterin pathway									
formaldehyde-activating enzyme		Fae	UNIPROT/Q9FA38	*Methylobacterium extorquens*	**DLM45_16005**	**formaldehyde-activating enzyme**	**79%**	**278**	**1e-98**
					DLM45_13410	aldehyde-activating enzyme	31%	80.5	7e-20
methylene-H_4MPT dehydrogenase A		MtdA	UNIPROT/P55818	*Methylobacterium extorquens*	DLM45_16030	methylene-H_4MPT dehydrogenase	30%	63.2	2e-12
NADP-dependent methylene-H_4MPT dehydrogenase B		MtdB	UNIPROT/O85012	*Methylobacterium extorquens*	**DLM45_16030**	**methylene-H_4MPT dehydrogenase**	**52%**	**279**	**4e-93**
methenyl-H_4MPT cyclohydrolase		Mch	UNIPROT/O85014	*Methylobacterium extorquens*	**DLM45_16020**	**methenyl-H_4MPT cyclohydrolase**	**56%**	**367**	**5e-127**
formyltransferase/ hydrolase complex	subunit A	FhcA	UNIPROT/C5B137	*Methylobacterium extorquens*	**DLM45_00145**	**formyl-MFR dehydrogenase subunit A**	**43%**	**426**	**3e-144**
	subunit B	FhcB	UNIPROT/C5B138	*Methylobacterium extorquens*	**DLM45_00140**	**formyl-MFR dehydrogenase**	**28%**	**78.2**	**2e-17**
	subunit C	FchC	UNIPROT/C5B135	*Methylosulfonomonas methylovora*	**DLM45_00155**	**Formyl-MFR dehydrogenase subunit C**	**43%**	**155**	**6e-47**
	Subunit D	FchD	UNIPROT/Q49118	*Methylosulfonomonas methylovora*	**DLM45_00150**	**Formyl-MFR H_4MPT N-formyltransferase**	**61%**	**363**	**2e-126**
(continued)									

Supplementary Table S5.4-6b: *BLAST search of the* H. methylovorum *strain Bras1 2018 draft for genes of formaldehyde metabolism (continued).*

FORMALDEHYDE METABOLISM *(continued)*								
REFERENCE SEQUENCE				BLAST RESULTS				
Enzyme	Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Formate oxidation								
formate dehydrogenase	Fdh	UNIPROT/P33160	*Pseudomonas* sp. 101	DLM45_10280	D-glycerate dehydrogenase	31%	120	8e-32
				DLM45_09195	D-glycerate dehydrogenase	34%	111	1e-28
Tetrahydrofolate pathway								
formate-THF ligase	Fhs	UNIPROT/Q83WS0	*Methylobacterium extorquens*	**DLM45_07050**	**formate H_4F ligase**	**68%**	**714**	**0.0**
				DLM45_13340	**formate H_4F ligase**	**65%**	**685**	**0.0**
bifunctional methenyl-H_4F cyclohydrolase/ methylene-H_4F dehydrogenase	FolD	UNIPROT/P24186	*Escherichia coli*	**DLM45_07540**	**bifunctional methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase, FolD**	**55%**	**273**	**4e-91**
				DLM45_13300	**bifunctional methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase, FolD**	**48%**	**253**	**3e-83**

Supplementary Table S5.4-7a: *BLAST search of the* H. denitrificans *strain ATCC 51888 genome for genes of inorganic sulfur compound metabolism.*

INORGANIC SULFUR METABOLISM								
REFERENCE SEQUENCE				BLAST RESULTS				
Enzyme	Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Sulfite metabolism								
sulfite reductase	CysJ	UNIPROT/P38038	*Escherichia coli*	**HDEN_RS07355**	**sulfite reductase subunit alpha**	**39%**	**380**	**2e-124**
SOE, molybdopterin oxidoreductase alpha subunit	SoeA	UNIPROT/D3RNN8	*Allochromatium vinosum*	HDEN_RS07070	molybdopterin oxidoreductase	26%	89.7	3e-19
				HDEN_RS14595	nitrate reductase	22%	72.4	4e-15
SOE, ferredoxin iron-sulfur binding domain protein	SoeB	UNIPROT/D3RNN7	*Allochromatium vinosum*	HDEN_RS04565	nitrate reductase subunit beta	36%	60.5	1e-11
SOE, DMSO reductase anchor subunit	SoeC	UNIPROT/D3RNN6	*Allochromatium vinosum*	HDEN_RS08105	DUF502 domain-containing protein	35%	35.0	0.004
Cysteine synthesis								
serine O-acetyltransferase	CysK/ MetA	UNIPROT/ A0A1D3PCK2	*Lactobacillus acidophilus*	HDEN_RS17400	cysteine synthase A, CysK	38%	155	4e-45
				HDEN_RS16975	cysteine synthase A, CysK	36%	147	8e-42
serine acetyltransferase	CysE	UNIPROT/P0A9D4	*Escherichia coli*	**HDEN_RS08820**	**serine O-acetyltransferase, CysE**	**51%**	**252**	**2e-83**
				HDEN_RS03690	serine acetyltransferase	36%	105	3e-27
				HDEN_RS00900	acetyltransferase	29%	35.0	0.003
cysteine synthase A	CysK	UNIPROT/P0ABK5	*Escherichia coli*	**HDEN_RS17400**	**cysteine synthase A, CysK**	**51%**	**238**	**2e-76**
				HDEN_RS16975	cysteine synthase A, CysK	36%	122	1e-32
cysteine synthase B	CysM	UNIPROT/P16703	*Escherichia coli*	HDEN_RS17400	cysteine synthase A, CysK	38%	155	4e-45
				HDEN_RS16975	cysteine synthase A, CysK	36%	147	8e-42
Homocysteine synthesis								
O-succinylhomoserine sulfhydrylase	MetZ	UNIPROT/P55218	*Pseudomonas aeruginosa*	**HDEN_RS00130**	**O-succinylhomoserine sulfhydrylase**	**46%**	**311**	**8e-103**
				HDEN_RS07495	O-acetylhomoserine sulfhydrylase/ O-succinylhomoserine sulfhydrylase	36%	244	1e-76
				HDEN_RS05360	cystathionine beta-lyase	33%	150	1e-41
Thiosulfate dehydrogenation								
thiosulfate dehydrogenase	TsdA	UNIPROT/D3RVD4	*Allochromatium vinosum*	HDEN_RS13585	cystathionine gamma-synthase	40%	147	2e-42
(continued)								

Supplementary Table S5.4-7b: *BLAST search of the* H. denitrificans *strain ATCC 51888 genome for genes of inorganic sulfur compound metabolism (continued).*

INORGANIC SULFUR METABOLISM *(continued)*									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Sox-mediated thiosulfate oxidation									
SOX complex	subunit A	SoxA	UNIPROT/O33434	*Paracoccus pantotrophus*	**HDEN_RS03420**	**sulfur oxidation c-type cytochrome SoxA**	**29%**	**90.9**	**3e-22**
			UNIPROT/D8JT38	*Hyphomicrobium denitrificans*	**HDEN_RS03420**	**sulfur oxidation c-type cytochrome SoxA**	**100%**	**581**	**0.0**
	subunit X	SoxX	UNIPROT/ A0A1I5IPQ3	*Paracoccus pantotrophus*	*no significant hits*				
			UNIPROT_D8JT37	*Hyphomicrobium denitrificans*	**HDEN_RS03415**	**sulfur oxidation c-type cytochrome SoxX**	**100%**	**415**	**7e-149**
	subunit B	SoxB	UNIPROT/ A0A1K2FBC2	*Paracoccus pantotrophus*	**HDEN_RS03435**	**thiosulfohydrolase SoxB**	**58%**	**646**	**0.0**
			UNIPROT/D8JTG4	*Hyphomicrobium denitrificans*	**HDEN_RS03435**	**thiosulfohydrolase SoxB**	**71%**	**1099**	**0.0**
	subunit C	SoxC	UNIPROT/A0A1I5IPK4	*Paracoccus pantotrophus*	**HDEN_RS05685**	**sulfite dehydrogenase**	**45%**	**313**	**8e-103**
			UNIPROT/D8JVS1	*Hyphomicrobium denitrificans*	**HDEN_RS05685**	**sulfite dehydrogenase**	**100%**	**813**	**0.0**
	subunit D	SoxD	UNIPROT/A0A1I5INZ2	*Paracoccus pantotrophus*	HDEN_RS05680	cytochrome c	36%	116	2e-31
					HDEN_RS07325	cytochrome c family protein	48%	90.9	5e-23
					HDEN_RS13510	cytochrome c family protein	42%	76.6	3e-17
					HDEN_RS10760	cytochrome c	35%	60.5	5e-12
	subunit Y	SoxY	UNIPROT/A0A1I5KIK2	*Paracoccus pantotrophus*	HDEN_RS05690	quinoprotein dehydrogenase-associated SoxYZ-like carrier	43%	179	3e-55
					HDEN_RS04930	quinoprotein dehydrogenase-associated SoxYZ-like carrier	42%	169	2e-51
					HDEN_RS01635	quinoprotein dehydrogenase-associated SoxYZ-like carrier	30%	53.5	2e-09
(continued)									

Supplementary Table S5.4-7c: *BLAST search of the* H. denitrificans *strain ATCC 51888 genome for genes of inorganic sulfur compound metabolism (continued).*

INORGANIC SULFUR METABOLISM *(continued)*									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Sox-mediated thiosulfate oxidation									
SOX complex (continued)	subunit Y	SoxY	UNIPROT/D8JT39	*Hyphomicrobium denitrificans*	**HDEN_RS03425**	**thiosulfate oxidation carrier protein SoxY**	**100%**	**267**	**6e-93**
					HDEN_RS06905	thiosulfate-binding protein SoxY	44%	102	3e-28
					HDEN_RS05690	quinoprotein dehydrogenase-associated SoxYZ-like carrier	29%	42.7	2e-06
	subunit Z	SoxZ	UNIPROT/ A0A089NU96	*Methylobacterium oryzae*	HDEN_RS05690	quinoprotein dehydrogenase-associated SoxYZ-like carrier	46%	201	2e-63
					HDEN_RS04930	quinoprotein dehydrogenase-associated SoxYZ-like carrier	41%	191	8e-60
					HDEN_RS01635	quinoprotein dehydrogenase-associated SoxYZ-like carrier	25%	65.9	1e-13
	subunit Z	SoxZ	UNIPROT/D8JX75	*Hyphomicrobium denitrificans*	**HDEN_RS06900**	**thiosulfate oxidation carrier complex protein SoxZ**	**100%**	**198**	**4e-67**
					HDEN_RS03430	thiosulfate oxidation carrier complex protein SoxZ	53%	111	5e-33
Hdr-mediated thiosulfate oxidation									
heterodisulfide reductase	**subunit A**	**HdrA**	**UNIPROT/D8JT26**	***Hyphomicrobium denitrificans***	**HDEN_RS03360**	**CoB--CoM heterodisulfide reductase iron-sulfur subunit A family protein**	**100%**	**673**	**0.0**
	subunit B	**HdrB**	**UNIPROT/ A0A088STN3**	***Acidithiobacillus caldus***	**HDEN_RS03375**	**disulfide reductase**	**23%**	**409**	**9e-144**
	subunit C	**HdrC**	**UNIPROT/ A0A088S946**	***Acidithiobacillus caldus***	**HDEN_RS03370**	**heterodisulfide reductase subunit C**	**52%**	**272**	**8e-92**
					HDEN_RS03350	4Fe-4S dicluster domain-containing protein	29%	39.7	7e-05
sulfur carrier protein		**TusA**	**UNIPROT/P0A890**	***Escherichia coli***	**HDEN_RS03395**	**sulfurtransferase TusA family protein**	**39%**	**50.4**	**2e-10**
lipoate binding protein		**LbpA**	**UNIPROT/D8JT31**	***Hyphomicrobium denitrificans***	**HDEN_RS03385**	**glycine cleavage system protein H**	**100%**	**299**	**8e-106**
					HDEN_RS06405	glycine cleavage system protein GcvH	28%	53.5	8e-11
(continued)									

Supplementary Table S5.4-7d: *BLAST search of the* H. denitrificans *strain ATCC 51888 genome for genes of inorganic sulfur compound metabolism (continued).*

INORGANIC SULFUR METABOLISM *(continued)*									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Sulfate adenylation pathway									
adenylyl-sulfate kinase		CysC	UNIPROT/P57702	*Pseudomonas aeruginosa*	*no significant hits*				
sulfate adenylyl-transferase	subunit 1	CysN	UNIPROT/P23845	*Escherichia coli*	HDEN_RS05165	adenylyl-sulfate kinase	45%	340	2e-111
					HDEN_RS05010	elongation factor Tu	45%	75.1	2e-15
	subunit 2	CysD	UNIPROT/O50273	*Pseudomonas aeruginosa*	**HDEN_RS05170**	**sulfate adenylyltransferase subunit CysD**	**63%**	**419**	**9e-148**
bifunctional sulfate adenylyl-transferase/ adenylyl-sulfate kinase	subunit 1	CysNC	UNIPROT/O50274	*Pseudomonas aeruginosa*	**HDEN_RS05165**	**adenylyl-sulfate kinase**	**43%**	**328**	**4e-105**
					HDEN_RS05010	elongation factor Tu	25%	75.9	2e-15
phosphoadenosine phosphosulfate reductase		CysH	UNIPROT/O05927	*Pseudomonas aeruginosa*	**HDEN_RS05175**	**phosphoadenylyl-sulfate reductase**	**39%**	**155**	**4e-46**
Sulfide oxidation									
sulfide:quinone oxidoreductase		SqrB	UNIPROT/D8JTH6	*Hyphomicrobium denitrificans*	**HDEN_RS03495**	**NAD(P)/FAD-dependent oxidoreductase**	**100%**	**814**	**0.0**
persulfide dioxygenase		Pdo	UNIPROT/D8JTH6	*Hyphomicrobium denitrificans*	**HDEN_RS03605**	**MBL fold metallo-hydrolase**	**100%**	**575**	**0.0**
rhodanase-type sulfur transferase		Rhd	UNIPROT/D8JTH7	*Hyphomicrobium denitrificans*	**HDEN_RS03500**	**TIGR01244 family phosphatase**	**100%**	**276**	**1e-96**
sulfide dehydrogenase	subunit A	FccA	UNIPROT/Q06529	*Allochromatium vinosum*	HDEN_RS04380	cytochrome c	27%	50.4	7e-09
	subunit B	FccB	UNIPROT/Q06530	*Allochromatium vinosum*	**HDEN_RS06910**	**flavocytochrome C sulfide dehydrogenase**	**43%**	**247**	**4e-77**

Supplementary Table S5.4-8a: *BLAST search of the* H. methylovorum *strain Bras1 2018 draft for genes of inorganic sulfur compound metabolism.*

INORGANIC SULFUR METABOLISM								
REFERENCE SEQUENCE				BLAST RESULTS				
Enzyme	Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Sulfite metabolism								
sulfite reductase	CysJ	UNIPROT/P38038	*Escherichia coli*	**DLM45_16080**	**sulfite reductase subunit alpha**	**37%**	**357**	**1e-115**
				DLM45_12320	oxidoreductase	43%	40.8	1e-04
SOE, molybdopterin oxidoreductase alpha subunit	SoeA	UNIPROT/D3RNN8	*Allochromatium vinosum*	DLM45_15805	molybdopterin oxidoreductase	26%	88.6	6e-19
				DLM45_12250	molybdopterin oxidoreductase	22%	72.4	8e-14
SOE, ferredoxin iron-sulfur binding domain protein	SoeB	UNIPROT/D3RNN7	*Allochromatium vinosum*	*no significant hits*				
SOE, DMSO reductase anchor subunit	SoeC	UNIPROT/D3RNN6	*Allochromatium vinosum*	*no significant hits*				
Cysteine synthesis								
serine O-acetyltransferase	CysK/ MetA	UNIPROT/ A0A1D3PCK2	*Lactobacillus acidophilus*	**DLM45_09210**	**cysteine synthase A, CysK**	**41%**	**158**	**5e-46**
				DLM45_04355	threonine ammonia-lyase, ilvA	26%	40.4	1e-04
serine acetyltransferase	CysE	UNIPROT/P0A9D4	*Escherichia coli*	**DLM45_01015**	**serine O-acetyltransferase, CysE**	**48%**	**243**	**1e-80**
				DLM45_04080	serine acetyltransferase	35%	97.4	2e-24
				DLM45_02620	UDP-3-O-(3-hydroxymyristoyl)glucosamine N-acyltransferas	28%	35.0	7e-04
cysteine synthase A	CysK	UNIPROT/P0ABK5	*Escherichia coli*	**DLM45_09210**	**cysteine synthase A, CysK**	**53%**	**246**	**9e-80**
cysteine synthase B	CysM	UNIPROT/P16703	*Escherichia coli*	DLM45_09210	cysteine synthase A, CysK	41%	158	5e-46
				DLM45_04355	threonine ammonia-lyase, ilvA	26%	40.5	1e-04
Homocysteine synthesis								
O-succinylhomoserine sulfhydrylase	MetZ	UNIPROT/P55218	*Pseudomonas aeruginosa*	**DLM45_09555**	**O-succinylhomoserine sulfhydrylase, MetZ**	**45%**	**311**	**1e-102**
				DLM45_00235	**O-acetylhomoserine sulfhydrylase/ O-succinylhomoserine sulfhydrylase**	**37%**	**248**	**3e-78**
				DLM45_00500	cystathionine beta-lyase, MetC	32%	149	4e-41
Thiosulfate dehydrogenation								
thiosulfate dehydrogenase	TsdA	UNIPROT/D3RVD4	*Allochromatium vinosum*	**DLM45_14295**	**hypothetical protein**	**37%**	**126**	**1e-34**
(continued)								

Supplementary Table S5.4-8b: *BLAST search of the* H. methylovorum *strain Bras1 2018 draft for genes of inorganic sulfur compound metabolism (continued).*

INORGANIC SULFUR METABOLISM *(continued)*									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Sox-mediated thiosulfate oxidation									
SOX complex	subunit A	SoxA	UNIPROT/O33434	*Paracoccus pantotrophus*	**DLM45_14680**	**sulfur oxidation c-type cytochrome SoxA**	**43%**	**220**	**1e-70**
					DLM45_12165	**sulfur oxidation protein SoxA**	**28%**	**93.2**	**4e-23**
			UNIPROT/D8JT38	*Hyphomicrobium denitrificans*	**DLM45_12165**	**sulfur oxidation c-type cytochrome SoxA**	**72%**	**401**	**5e-142**
					DLM45_14680	**sulfur oxidation c-type cytochrome SoxA**	**26%**	**95.1**	**1e-23**
	subunit X	SoxX	UNIPROT/ A0A1I5IPQ3	*Paracoccus pantotrophus*	**DLM45_14665**	**sulfur oxidation protein SoxX**	**40%**	**109**	**4e-31**
					DLM45_12160	**sulfur oxidation c-type cytochrome SoxX**	**30%**	**33.9**	**0.003**
			UNIPROT_D8JT37	*Hyphomicrobium denitrificans*	**DLM45_12160**	**sulfur oxidation c-type cytochrome SoxX**	**54%**	**216**	**4e-71**
	subunit B	SoxB	UNIPROT/ A0A1K2FBC2	*Paracoccus pantotrophus*	**DLM45_14685**	**thiosulfohydrolase SoxB**	**62%**	**704**	**0.0**
					DLM45_12185	**thiosulfohydrolase SoxB**	**62%**	**691**	**0.0**
			UNIPROT/D8JTG4	*Hyphomicrobium denitrificans*	**DLM45_12185**	**thiosulfohydrolase SoxB**	**71%**	**823**	**0.0**
					DLM45_14685	**thiosulfohydrolase SoxB**	**60%**	**691**	**0.0**
	subunit C	SoxC	UNIPROT/A0A1I5IPK4	*Paracoccus pantotrophus*	**DLM45_14690**	**sulfite dehydrogenase soxC**	**66%**	**546**	**0.0**
					DLM45_12190	**sulfite dehydrogenase soxC**	**62%**	**511**	**5e-180**
					DLM45_02965	**sulfite dehydrogenase soxC**	**43%**	**310**	**6e-102**
			UNIPROT/D8JVS1	*Hyphomicrobium denitrificans*	**DLM45_02965**	**sulfite dehydrogenase soxC**	**49%**	**356**	**1e-119**
					DLM45_12190	**sulfite dehydrogenase soxC**	**46%**	**339**	**9e-113**
					DLM45_14690	**sulfite dehydrogenase soxC**	**46%**	**323**	**1e-106**
	subunit D	SoxD	UNIPROT/A0A1I5INZ2	*Paracoccus pantotrophus*	**DLM45_12195**	**cytochrome c**	**55%**	**223**	**4e-71**
					DLM45_14695	**cytochrome c**	**52%**	**204**	**1e-63**
					DLM45_02970	**cytochrome c**	**38%**	**122**	**5e-34**
					DLM45_16050	cytochrome c family protein	47%	92.4	2e-23
(continued)									

Supplementary Table S5.4-8c: *BLAST search of the* H. methylovorum *strain Bras1 2018 draft for genes of inorganic sulfur compound metabolism (continued).*

INORGANIC SULFUR METABOLISM *(continued)*									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Sox-mediated thiosulfate oxidation									
SOX complex (continued)	subunit Y	SoxY	UNIPROT/A0A1I5KIK2	*Paracoccus pantotrophus*	DLM45_01870	quinoprotein dehydrogenase-associated SoxYZ-like carrier	40%	152	2e-45
					DLM45_03680	quinoprotein dehydrogenase-associated SoxYZ-like carrier	41%	147	2e-43
					DLM45_12175	thiosulfate oxidation carrier protein SoxY	28%	42.7	3e-06
					DLM45_14670	thiosulfate oxidation carrier protein SoxY	34%	42.4	4e-06
					DLM45_01255	thiosulfate oxidation carrier protein SoxY	32%	42.0	8e-06
	subunit Y	SoxY	UNIPROT/D8JT39	*Hyphomicrobium denitrificans*	**DLM45_12175**	**thiosulfate oxidation carrier protein SoxY**	**79%**	**201**	**8e-67**
					DLM45_14670	**thiosulfate oxidation carrier protein SoxY**	**54%**	**112**	**3e-32**
					DLM45_01255	**thiosulfate oxidation carrier protein SoxY**	**47%**	**107**	**3e-30**
	subunit Z	SoxZ	UNIPROT/ A0A089NU96	*Methylobacterium oryzae*	DLM45_01870	quinoprotein dehydrogenase-associated SoxYZ-like carrier	42%	187	1e-58
					DLM45_03680	quinoprotein dehydrogenase-associated SoxYZ-like carrier	38%	166	2e-50
	subunit Z	SoxZ	UNIPROT/D8JX75	*Hyphomicrobium denitrificans*	**DLM45_01260**	**thiosulfate oxidation carrier complex protein SoxZ**	**77%**	**155**	**3e-50**
					DLM45_12180	**sulfur oxidation protein soxZ**	**53%**	**108**	**7e-32**
					DLM45_14675	**sulfur oxidation protein soxZ**	**51%**	**102**	**1e-29**
(continued)									

Supplementary Table S5.4-8d: *BLAST search of the* H. methylovorum *strain Bras1 2018 draft for genes of inorganic sulfur compound metabolism (continued).*

INORGANIC SULFUR METABOLISM *(continued)*									
REFERENCE SEQUENCE					BLAST RESULTS				
Enzyme		Name	Database/Accession	Origin species	Gene ID	NCBI annotation	Identity	Score (Bits)	E value
Hdr-mediated thiosulfate oxidation									
heterodisulfide reductase	subunit A	HdrA	UNIPROT/D8JT26	*Hyphomicrobium denitrificans*	*no significant hits*				
	subunit B	HdrB	UNIPROT/ A0A088STN3	*Acidithiobacillus caldus*	*no significant hits*				
	subunit C	HdrC	UNIPROT/ A0A088S946	*Acidithiobacillus caldus*	DLM45_11655	glycolate oxidase iron-sulfur subunit	28%	40.8	4e-05
sulfur carrier protein		TusA	UNIPROT/P0A890	*Escherichia coli*	*no significant hits*				
lipoate binding protein		LbpA	UNIPROT/D8JT31	*Hyphomicrobium denitrificans*	*no significant hits*				
Sulfate adenylation pathway									
adenylyl-sulfate kinase		CysC	UNIPROT/P57702	*Pseudomonas aeruginosa*	*no significant hits*				
sulfate adenylyl-transferase	subunit 1	CysN	UNIPROT/P23845	*Escherichia coli*	**DLM45_01935**	**sulfate adenylyltransferase**	**43%**	**341**	**1e-112**
					DLM45_02090	elongation factor Tu	27%	74.7	5e-16
	subunit 2	CysD	UNIPROT/O50273	*Pseudomonas aeruginosa*	**DLM45_01930**	**sulfate adenylyltransferase subunit CysD**	**62%**	**411**	**9e-145**
bifunctional sulfate adenylyl-transferase/ adenylyl-sulfate kinase	subunit 1	CysNC	UNIPROT/O50274	*Pseudomonas aeruginosa*	DLM45_01935	sulfate adenylyltransferase	43%	332	3e-107
					DLM45_02090	elongation factor Tu	26%	75.9	3e-16
phosphoadenosine phosphosulfate reductase		CysH	UNIPROT/O05927	*Pseudomonas aeruginosa*	DLM45_01925	phosphoadenylyl-sulfate reductase	39%	164	5e-50

Table S5.4-8e: *BLAST search of the* H. methylovorum *strain Bras1 2018 draft for genes of inorganic sulfur compound metabolism (continued).*

Sulfide oxidation									
sulfide:quinone oxidoreductase		SqrB	UNIPROT/D8JTH6	*Hyphomicrobium denitrificans*	DLM45_01250	flavocytochrome C	21%	58.5	6e-11
Persulfide dioxygenase		Pdo	UNIPROT/D8JTH6	*Hyphomicrobium denitrificans*	*no significant hits*				
rhodanese-type sulfur transferase		Rhd	UNIPROT/D8JTH7	*Hyphomicrobium denitrificans*	*no significant hits*				
sulfide dehydrogenase	subunit A	FccA	UNIPROT/Q06529	*Allochromatium vinosum*	*no significant hits*				
	subunit B	FccB	UNIPROT/Q06530	*Allochromatium vinosum*	**DLM45_01250**	**flavocytochrome C**	**43%**	**243**	**2e-76**

Supplementary Table S5.5-1a: *Comparative proteomics of* H. denitrificans*, proteins of interest intensity data.* Unidentified proteins are marked as N/A.

Enzyme				Unique Peptides						MS/MS count					
Protein ID	Protein Name	Gene Locus	Pathway	Control			Carbon			Control			Carbon		
				1	2	3	1	2	3	1	2	3	1	2	3
HDEN_RS00130	MetZ	28355:29563 Reverse	S	25	23	22	22	21	18	34	23	30	26	28	20
HDEN_RS00515	FolD	122735:123637 Reverse	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS00520	Fhs	123720:125396 Reverse	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS01635	SoxYZ	345962:346750 Forward	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS02580	catalase	551292:552272 Forward	MSC	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS02970	Fdh	638207:639406 Reverse	HCHO	23	17	23	22	24	23	29	15	40	30	35	31
HDEN_RS03360	HdrA	716665:717714 Forward	S	1	1	2	19	19	19	0	0	1	27	28	32
HDEN_RS03370	HdrC	718580:719362 Forward	S	1	0	0	16	16	17	0	0	0	22	24	22
HDEN_RS03375	HdrB	719411:720355 Forward	S	0	0	0	15	14	14	0	0	0	27	24	28
HDEN_RS03385	LbpA	720866:721306 Forward	S	0	0	0	2	3	3	0	0	0	9	9	11
HDEN_RS03395	TusA	722917:723198 Reverse	S	0	0	0	4	3	4	0	0	0	6	4	5
HDEN_RS03415	SoxX	725184:725852 Reverse	S	0	1	0	4	6	4	0	0	0	5	7	6
HDEN_RS03420	SoxA	725830:726663 Reverse	S	3	2	1	17	17	18	0	0	0	33	24	30
HDEN_RS03425	SoxY	726815:727276 Forward	S	0	0	0	3	3	3	0	0	0	8	8	10
HDEN_RS03430	SoxZ	727314:727643 Forward	S	0	0	0	6	6	6	0	0	0	15	16	17
HDEN_RS03435	SoxB	727686:729374 Forward	S	3	2	3	52	51	52	0	0	1	124	111	111
HDEN_RS03495	SqrB	744049:745296 Forward	S	0	0	1	27	25	24	0	0	0	54	39	38
HDEN_RS03500	Rhd	745472:745921 Forward	S	0	0	1	11	10	10	0	0	0	20	20	17
HDEN_RS03605	Pdo	764123:765016 Forward	S	6	3	9	14	16	16	2	1	4	22	28	22
HDEN_RS03620	MtoX	766986:768272 Reverse	MSC	14	12	14	52	51	51	11	9	12	505	465	478
HDEN_RS04930	SoxYZ	1054621:1055433 Reverse	S	8	4	7	10	8	7	4	1	4	14	7	8
HDEN_RS05165	CysNC	1094436:1095968 Reverse	S	17	12	20	1	1	2	19	8	21	1	0	2
HDEN_RS05170	CysD	1095968:1096873 Reverse	S	7	3	10	0	0	0	9	3	17	0	0	0

Supplementary Table S5.5-1b: *Comparative proteomics of* H. denitrificans*, proteins of interest intensity data (continued).* Unidentified proteins are marked as N/A.

Enzyme	Carbon versus Control			Intensity						LFQ Intensity					
Protein ID	p-value	Welch's q-value	Fold change	Control			Carbon			Control			Carbon		
				1	2	3	1	2	3	1	2	3	1	2	3
HDEN_RS00130	0.010	0.023	0.62	1.25E+09	1.31E+09	1.34E+09	3.35E+09	1.12E+08	2.25E+09	31.17	30.70	31.10	30.34	30.33	30.20
HDEN_RS00515	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS00520	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS01635	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS02580	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS02970	0.140	0.150	1.21	1.06E+09	9.39E+08	9.80E+08	1.52E+09	2.46E+07	8.71E+08	29.56	29.21	29.67	29.66	29.77	29.82
HDEN_RS03360	0.000	0.000	124.12	2.89E+09	2.53E+09	2.14E+09	2.44E+07	9.13E+05	6.03E+06	24.22	24.16	23.91	30.86	31.17	31.13
HDEN_RS03370	0.000	0.000	41.84	1.52E+09	1.29E+09	8.58E+08	0.00	0.00	1.23E+06	24.97	24.98	23.96	29.55	30.23	30.29
HDEN_RS03375	0.000	0.004	64.61	1.46E+09	1.22E+09	1.41E+09	0.00	0.00	0.00	24.48	23.69	24.40	30.21	30.21	30.20
HDEN_RS03385	0.000	0.003	21.58	4.76E+08	4.36E+08	4.82E+08	0.00	0.00	0.00	23.82	24.63	24.11	28.56	28.70	28.60
HDEN_RS03395	0.000	0.006	6.34	9.67E+07	1.16E+08	1.27E+08	0.00	0.00	0.00	23.73	24.46	23.75	26.50	26.74	26.71
HDEN_RS03415	0.000	0.002	12.54	2.86E+08	2.94E+08	2.74E+08	0.00	1.28E+06	0.00	24.64	23.83	24.43	27.71	27.94	28.21
HDEN_RS03420	0.000	0.000	53.70	1.67E+09	1.26E+09	1.44E+09	4.23E+06	4.21E+05	3.59E+07	24.05	24.86	24.56	30.84	29.83	30.05
HDEN_RS03425	0.000	0.000	39.65	1.31E+09	1.14E+09	1.19E+09	0.00	0.00	0.00	24.62	24.39	25.11	30.12	30.08	29.84
HDEN_RS03430	0.000	0.005	65.72	1.63E+09	1.39E+09	1.42E+09	0.00	0.00	0.00	24.61	23.81	24.29	30.23	30.29	30.30
HDEN_RS03435	0.000	0.002	5000.31	1.33E+10	1.07E+10	1.34E+10	1.42E+07	3.41E+05	2.27E+07	21.81	20.64	20.77	33.51	33.25	33.33
HDEN_RS03495	0.000	0.004	102.77	1.77E+09	1.54E+09	4.76E+09	5.40E+06	0.00	0.00	24.43	24.55	23.88	31.95	30.48	30.48
HDEN_RS03500	0.000	0.007	89.03	1.86E+09	1.39E+09	1.63E+09	9.46E+06	0.00	0.00	23.89	24.76	23.23	30.51	30.28	30.51
HDEN_RS03605	0.002	0.011	39.37	1.69E+09	1.36E+09	1.41E+09	1.27E+08	1.10E+06	3.61E+07	25.15	23.78	26.15	30.19	30.34	30.44
HDEN_RS03620	0.000	0.000	570.55	2.22E+11	1.97E+11	2.47E+11	4.50E+08	2.10E+07	2.77E+08	27.94	28.86	28.06	37.62	37.38	37.34
HDEN_RS04930	0.014	0.012	1.79	2.45E+08	2.67E+08	4.32E+08	1.94E+08	3.57E+06	1.24E+08	27.07	27.16	26.76	28.14	27.79	27.59
HDEN_RS05165	0.000	0.000	0.03	3.75E+07	1.60E+07	1.67E+07	7.68E+08	1.25E+07	4.24E+08	28.70	28.25	28.58	23.34	24.04	23.57
HDEN_RS05170	0.005	0.011	0.13	0.00	0.00	0.00	3.58E+08	3.33E+06	1.42E+08	26.64	27.35	27.19	24.82	23.19	24.26

Supplementary Table S5.5-1c: *Comparative proteomics of* H. denitrificans*, proteins of interest intensity data (continued).* Unidentified proteins are marked as N/A.

Enzyme				Unique Peptides						MS/MS count					
Protein ID	Protein Name	Gene Locus	Pathway	Control			Carbon			Control			Carbon		
				1	2	3	1	2	3	1	2	3	1	2	3
HDEN_RS05175	CysH	1096870:1097679 Reverse	S	4	3	5	0	0	0	3	3	7	0	0	0
HDEN_RS05680	SoxD	1195731:1196273 Reverse	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS05685	SoxC	1196260:1197558 Reverse	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS05690	SoxYZ	1197577:1198392 Reverse	S	0	0	3	8	6	7	0	0	0	12	6	5
HDEN_RS06435	XoxF	1331041:1332867 Reverse	MeOH	22	20	24	23	20	24	29	25	56	27	26	33
HDEN_RS06510	MxaJ	1344061:1344366 Reverse	MeOH	7	5	7	7	8	5	10	5	11	8	10	5
HDEN_RS06515	MxaI	1344467:1345066 Reverse	MeOH	7	4	7	5	5	6	15	7	17	9	9	9
HDEN_RS06520	MxaG	1345107:1346018 Reverse	MeOH	26	23	24	27	24	25	56	34	39	71	56	40
HDEN_RS06525	MxaF	1346250:1348139 Reverse	MeOH	66	54	64	65	66	63	650	595	588	581	545	492
HDEN_RS06900	SoxZ	1449412:1449735 Reverse	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS06905	SoxY	1449746:1450231 Reverse	S	3	1	2	3	2	3	3	0	3	4	3	5
HDEN_RS07280	Fae	1526339:1526851 Reverse	HCHO	16	14	16	17	16	17	112	78	147	109	103	106
HDEN_RS07295	Mch	1528906:1529871 Reverse	HCHO	15	11	16	14	14	16	30	15	34	21	16	28
HDEN_RS07305	MtdB	1531029:1531931 Reverse	HCHO	22	19	24	24	23	24	60	48	82	54	66	62
HDEN_RS07325	SoxD	1534075:1534431 Reverse	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS07355	CysJ	1538660:1540447 Forward	S	19	21	22	0	0	1	22	14	25	0	0	0
HDEN_RS07495	MetZ	1569550:1570839 Forward	S	21	18	23	26	25	24	29	19	30	49	46	33
HDEN_RS07890	FhcC	1629123:1629932 Reverse	HCHO	20	16	19	18	19	20	19	17	25	26	30	27
HDEN_RS07895	FhcD	1629932:1630834 Reverse	HCHO	11	12	13	13	13	13	20	23	27	25	19	26
HDEN_RS07900	FhcA	1630831:1632495 Reverse	HCHO	36	31	38	39	36	35	67	44	76	72	58	75
HDEN_RS07905	FhcB	1632516:1633802 Reverse	HCHO	25	21	25	24	25	25	52	40	43	46	46	53
HDEN_RS07935	XoxF	1638341:1640167 Forward	MeOH	21	21	26	23	22	23	62	62	145	78	63	111
HDEN_RS08820	CysE	1820068:1820895 Reverse	S	12	9	12	12	10	9	8	8	14	9	9	7

Supplementary Table S5.5-1d: *Comparative proteomics of* H. denitrificans*, proteins of interest intensity data (continued).* Unidentified proteins are marked as N/A.

Enzyme	Carbon versus Control			Intensity						LFQ Intensity					
Protein ID	p-value	Welch's q-value	Fold change	Control			Carbon			Control			Carbon		
				1	2	3	1	2	3	1	2	3	1	2	3
HDEN_RS05175	0.007	0.024	0.27	0.00	0.00	0.00	1.70E+08	3.10E+06	7.80E+07	26.14	26.32	26.22	24.44	24.92	23.65
HDEN_RS05680	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS05685	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS05690	0.008	0.008	5.42	1.68E+08	1.30E+08	1.84E+08	2.63E+07	0.00	0.00	23.87	25.04	24.40	27.48	26.26	26.87
HDEN_RS06435	0.818	0.899	0.90	2.24E+09	1.08E+09	1.07E+09	6.53E+09	5.66E+07	8.32E+08	29.57	30.22	31.54	29.95	30.07	30.83
HDEN_RS06510	0.631	0.740	0.84	4.83E+08	5.41E+08	5.22E+08	8.56E+08	3.26E+07	8.91E+08	30.03	28.74	28.79	29.12	29.24	28.42
HDEN_RS06515	0.221	0.197	0.76	4.69E+08	5.26E+08	5.90E+08	1.17E+09	2.09E+07	8.94E+08	29.65	28.87	29.29	29.00	29.06	28.58
HDEN_RS06520	0.080	0.069	1.50	5.28E+09	5.34E+09	8.01E+09	6.81E+09	1.30E+08	4.08E+09	31.90	31.45	31.86	32.72	32.24	32.02
HDEN_RS06525	0.646	0.752	0.93	2.19E+11	2.13E+11	2.80E+11	3.35E+11	1.27E+10	2.44E+11	37.84	37.48	37.39	37.76	37.41	37.24
HDEN_RS06900	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS06905	0.001	0.004	2.58	7.91E+07	5.05E+07	6.91E+07	3.33E+07	2.93E+05	2.66E+07	24.45	24.22	24.69	25.69	26.00	25.77
HDEN_RS07280	0.057	0.065	0.77	4.11E+10	3.41E+10	3.31E+10	6.59E+10	2.25E+09	3.29E+10	34.85	35.30	35.07	34.59	34.70	34.79
HDEN_RS07295	0.003	0.006	0.79	1.64E+09	1.22E+09	1.28E+09	2.62E+09	6.14E+07	1.50E+09	30.36	30.49	30.47	30.08	30.16	30.08
HDEN_RS07305	0.175	0.143	0.89	5.82E+09	5.50E+09	4.56E+09	1.07E+10	2.98E+08	4.59E+09	32.21	32.25	32.46	32.04	32.28	32.08
HDEN_RS07325	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS07355	0.006	0.022	0.04	5.56E+06	0.00	0.00	8.56E+08	2.30E+07	5.59E+08	29.17	28.81	28.89	24.61	22.45	25.47
HDEN_RS07495	0.000	0.002	2.42	3.28E+09	3.25E+09	3.52E+09	2.10E+09	5.47E+07	1.06E+09	30.27	30.14	30.17	31.59	31.48	31.33
HDEN_RS07890	0.064	0.088	1.28	1.48E+09	1.27E+09	1.24E+09	1.80E+09	4.27E+07	1.06E+09	29.94	29.55	29.98	30.12	30.24	30.18
HDEN_RS07895	0.102	0.087	1.12	2.38E+09	2.23E+09	2.15E+09	3.08E+09	1.05E+08	1.46E+09	30.70	30.89	30.68	30.84	30.95	30.99
HDEN_RS07900	0.163	0.174	1.18	5.40E+09	4.26E+09	5.38E+09	7.22E+09	1.35E+08	4.20E+09	31.89	31.50	31.89	32.04	31.91	32.04
HDEN_RS07905	0.210	0.229	1.12	5.03E+09	4.44E+09	5.11E+09	6.46E+09	1.25E+08	4.15E+09	31.89	31.57	31.86	31.99	31.92	31.88
HDEN_RS07935	0.701	0.812	0.83	1.22E+10	4.98E+09	5.61E+09	3.00E+10	2.75E+08	4.31E+09	32.06	32.34	33.91	32.22	32.11	33.16
HDEN_RS08820	0.014	0.011	0.80	2.11E+08	1.82E+08	2.16E+08	4.04E+08	8.87E+06	2.59E+08	27.74	27.66	27.58	27.23	27.34	27.44

Supplementary Table S5.5-1e: *Comparative proteomics of* H. denitrificans*, proteins of interest intensity data (continued).* Unidentified proteins are marked as N/A.

Enzyme				Unique Peptides						MS/MS count					
Protein ID	Protein Name	Gene Locus	Pathway	Control			Carbon			Control			Carbon		
				1	2	3	1	2	3	1	2	3	1	2	3
HDEN_RS09180	XoxF	1900170:1901993 Reverse	MeOH	13	7	12	12	10	10	11	6	12	12	12	9
HDEN_RS09220	Fae	1908727:1909266 Forward	HCHO	10	8	8	7	8	7	15	8	21	11	13	8
HDEN_RS10485	Fae	2180153:2180710 Forward	HCHO	6	5	9	6	5	7	4	1	10	7	5	3
HDEN_RS10760	SoxD	2238912:2239283 Forward	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS12420	catalase	2571897:2573372 Reverse	MSC	31	26	35	43	40	41	42	31	45	104	60	69
HDEN_RS12910	catalase	2678902:2679882 Forward	MSC	9	5	10	11	9	10	6	2	12	15	8	10
HDEN_RS13510	SoxD	2809698:2810267 Forward	S	8	4	7	3	7	8	7	2	9	3	3	3
HDEN_RS15500	Fhs	3201592:3203268 Reverse	HCHO	7	8	6	6	6	6	81	48	82	88	92	91
HDEN_RS15885	FolD	3282716:3283603 Forward	HCHO	24	19	25	21	24	24	55	43	63	40	45	49
HDEN_RS16975	CysK	3496108:3497145 Forward	S	10	6	10	24	22	23	6	1	5	46	49	50
HDEN_RS17110	Acs	3527571:3529520 Forward	S	42	42	42	39	37	36	80	69	73	66	58	58
HDEN_RS17400	CysK	3588958:3589959 Reverse	S	19	18	20	19	17	17	47	35	45	36	37	37

Supplementary Table S5.5-1f: *Comparative proteomics of* H. denitrificans*, proteins of interest intensity data (continued).* Unidentified proteins are marked as N/A.

Enzyme	Carbon versus Control			Intensity						LFQ Intensity					
Protein ID	p-value	Welch's q-value	Fold change	Control			Carbon			Control			Carbon		
				1	2	3	1	2	3	1	2	3	1	2	3
HDEN_RS09180	0.173	0.178	1.14	1.98E+08	1.89E+08	2.28E+08	2.98E+08	6.49E+06	2.01E+08	27.28	27.21	27.35	27.63	27.27	27.48
HDEN_RS09220	0.000	0.000	0.43	8.41E+08	8.80E+08	6.83E+08	3.13E+09	7.38E+07	1.52E+09	30.61	30.58	30.71	29.36	29.55	29.38
HDEN_RS10485	0.252	0.228	0.85	9.98E+07	8.44E+07	9.91E+07	2.95E+08	6.73E+06	9.07E+07	26.64	26.97	26.99	26.75	26.78	26.35
HDEN_RS10760	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HDEN_RS12420	0.002	0.013	3.60	8.87E+09	6.85E+09	1.20E+10	3.90E+09	8.49E+07	2.12E+09	31.00	30.98	31.04	33.40	32.54	32.63
HDEN_RS12910	0.004	0.007	1.59	3.41E+08	3.40E+08	3.97E+08	3.07E+08	4.41E+06	1.65E+08	27.34	27.33	27.53	28.24	28.01	27.95
HDEN_RS13510	0.042	0.028	0.71	1.70E+08	2.01E+08	8.66E+07	4.12E+08	4.67E+06	2.28E+08	27.54	27.33	27.76	26.82	27.22	27.10
HDEN_RS15500	0.105	0.117	1.11	8.78E+09	6.97E+09	7.90E+09	1.03E+10	2.60E+08	6.17E+09	32.58	32.35	32.51	32.64	32.59	32.67
HDEN_RS15885	0.009	0.024	0.57	3.53E+09	3.15E+09	3.15E+09	1.10E+10	2.22E+08	5.71E+09	32.38	32.03	32.60	31.46	31.57	31.54
HDEN_RS16975	0.000	0.007	33.68	4.25E+09	3.67E+09	3.94E+09	1.53E+08	2.23E+06	1.04E+08	27.12	25.91	26.94	31.72	31.76	31.71
HDEN_RS17110	0.003	0.013	0.62	4.81E+09	4.03E+09	4.74E+09	1.23E+10	2.92E+08	7.37E+09	32.79	32.45	32.72	32.01	31.93	31.97
HDEN_RS17400	0.093	0.081	0.90	2.63E+09	2.31E+09	2.76E+09	4.56E+09	1.03E+08	2.74E+09	31.24	31.09	31.28	31.11	31.08	30.99

Supplementary Table S5.5-2aa: *Comparative proteomics of* H. methylovorum*, proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			Unique Peptides									MS/MS count								
Protein ID	Protein Name	Pathway	MeOH			$DMSO_2$			DMS			MeOH			$DMSO_2$			DMS		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
DLM45_00110	XoxF	MeOH	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_00140	FchB	HCHO	18	20	18	14	16	15	24	24	23	23	24	22	14	13	15	40	45	42
DLM45_00145	FhcA	HCHO	28	27	31	23	21	22	32	33	32	37	42	44	18	19	26	60	66	70
DLM45_00150	FhcD	HCHO	14	14	14	11	11	10	15	16	16	17	16	11	8	10	5	22	26	34
DLM45_00155	FhcC	HCHO	20	18	16	12	15	16	21	20	22	21	18	13	2	7	7	19	24	29
DLM45_00235	MetZ	S	24	24	25	20	20	22	27	28	28	30	25	27	17	13	21	56	50	48
DLM45_01015	CysE	S	9	9	10	4	8	8	11	11	12	3	3	2	0	3	2	14	11	19
DLM45_01255	SoxY	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_01260	SoxZ	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_01870	SoxYZ	S	5	6	3	1	2	2	11	11	11	0	3	1	0	0	0	16	14	12
DLM45_01925	CysH	S	6	4	4	0	0	0	0	0	0	6	4	4	0	0	0	0	0	0
DLM45_01930	CysD	S	15	14	15	4	4	3	0	1	1	19	18	15	2	1	1	0	0	0
DLM45_01935	CysNC	S	22	24	23	2	1	1	1	1	1	23	25	24	0	0	0	0	0	0
DLM45_02930	DmoA	MSC	3	2	1	32	33	34	6	7	6	0	0	0	85	82	101	4	4	2
DLM45_02940	SfnF	MSC	1	1	1	3	3	4	0	0	0	0	1	0	2	2	3	0	0	0
DLM45_02945	SfnG	MSC	6	5	2	8	8	10	9	10	8	3	3	1	8	10	11	8	6	7
DLM45_02965	SoxC	S	6	5	6	16	15	15	7	7	6	2	3	2	17	13	16	2	4	2
DLM45_02970	SoxC	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_03560	MxaF	MeOH	55	52	55	44	43	45	56	56	56	282	255	245	155	164	162	359	394	368
DLM45_03565	MxaJ	MeOH	20	17	17	6	11	8	20	21	21	19	13	11	0	1	1	23	29	24
DLM45_03570	MxaG	MeOH	6	5	6	1	1	2	7	8	8	5	4	3	1	1	2	8	13	7
DLM45_03575	MxaI	MeOH	3	1	4	0	0	1	4	5	4	3	0	5	0	0	0	8	8	4
DLM45_03680	SoxYZ	S	2	3	2	0	0	1	4	5	6	1	2	0	0	0	0	3	2	5

Supplementary Table S5.5-2ab: *Comparative proteomics of* H. methylovorum, *proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			Intensity								
Protein ID	Protein Name	Pathway	MeOH			$DMSO_2$			DMS		
			1	2	3	1	2	3	1	2	3
DLM45_00110	XoxF	MeOH	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_00140	FchB	HCHO	4.77E+09	4.79E+09	3.42E+09	4.29E+08	5.42E+08	6.35E+08	6.45E+09	6.97E+09	9.16E+09
DLM45_00145	FhcA	HCHO	4.97E+09	4.50E+09	4.27E+09	1.89E+09	1.84E+09	1.93E+09	3.72E+09	5.51E+09	5.52E+09
DLM45_00150	FhcD	HCHO	3.81E+09	3.04E+09	1.85E+09	4.69E+08	7.23E+08	4.17E+08	2.95E+09	3.93E+09	4.99E+09
DLM45_00155	FhcC	HCHO	3.78E+09	2.28E+09	1.67E+09	2.41E+08	5.91E+08	6.83E+08	1.52E+09	2.35E+09	2.84E+09
DLM45_00235	MetZ	S	5.85E+09	5.51E+09	5.53E+09	1.41E+09	1.31E+09	1.44E+09	8.13E+09	7.17E+09	9.03E+09
DLM45_01015	CysE	S	2.17E+08	1.38E+08	1.19E+08	1.86E+07	6.53E+07	4.71E+07	4.03E+08	6.20E+08	6.20E+08
DLM45_01255	SoxY	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_01260	SoxZ	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_01870	SoxYZ	S	2.29E+08	2.16E+08	5.43E+07	7.27E+06	1.24E+07	7.91E+06	5.89E+08	5.98E+08	7.15E+08
DLM45_01925	CysH	S	2.81E+08	2.15E+08	1.19E+08	0.00	0.00	0.00	0.00	0.00	0.00
DLM45_01930	CysD	S	1.64E+09	1.54E+09	1.16E+09	5.65E+07	5.33E+07	5.60E+07	0.00	4.37E+06	6.21E+06
DLM45_01935	CysNC	S	1.98E+09	1.87E+09	1.51E+09	2.18E+07	1.96E+07	1.57E+07	4.14E+06	1.94E+06	3.05E+06
DLM45_02930	DmoA	MSC	5.28E+07	8.50E+07	3.79E+04	2.98E+10	3.07E+10	2.98E+10	1.33E+08	9.06E+07	5.47E+07
DLM45_02940	SfnF	MSC	1.32E+08	1.06E+08	8.70E+07	1.48E+08	2.00E+08	1.87E+08	0.00	0.00	0.00
DLM45_02945	SfnG	MSC	1.60E+08	1.14E+08	7.28E+07	1.33E+09	1.13E+09	1.30E+09	1.79E+08	1.50E+08	1.65E+08
DLM45_02965	SoxC	S	1.13E+08	8.88E+07	7.92E+07	5.21E+08	4.49E+08	4.94E+08	5.73E+07	9.71E+07	7.88E+07
DLM45_02970	SoxC	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_03560	MxaF	MeOH	2.21E+11	1.71E+11	1.45E+11	3.70E+10	3.95E+10	3.98E+10	2.37E+11	2.62E+11	2.33E+11
DLM45_03565	MxaJ	MeOH	2.33E+09	1.21E+09	5.80E+08	2.98E+07	1.54E+08	7.24E+07	1.18E+09	2.19E+09	2.18E+09
DLM45_03570	MxaG	MeOH	6.30E+08	4.60E+08	4.16E+08	1.35E+07	1.40E+07	2.72E+07	5.04E+08	7.87E+08	6.96E+08
DLM45_03575	MxaI	MeOH	2.23E+08	3.41E+07	1.97E+08	0.00	0.00	1.50E+07	3.26E+08	3.87E+08	1.89E+08
DLM45_03680	SoxYZ	S	4.52E+07	4.19E+07	2.01E+07	0.00	0.00	4.60E+06	4.75E+07	6.25E+07	1.15E+08

Supplementary Table S5.5-2ac: *Comparative proteomics of* H. methylovorum, *proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			LFQ Intensity								
Protein ID	Protein Name	Pathway	MeOH			$DMSO_2$			DMS		
			1	2	3	1	2	3	1	2	3
DLM45_00110	XoxF	MeOH	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_00140	FchB	HCHO	32.07	32.35	32.18	32.11	31.92	32.01	32.31	32.44	32.48
DLM45_00145	FhcA	HCHO	32.30	32.59	32.41	32.43	32.48	32.58	32.29	32.70	32.51
DLM45_00150	FhcD	HCHO	31.86	31.80	31.43	31.04	31.49	31.56	31.62	31.81	31.97
DLM45_00155	FhcC	HCHO	31.55	31.42	30.88	30.11	31.28	31.27	30.91	31.63	31.64
DLM45_00235	MetZ	S	32.40	32.49	32.74	32.87	32.70	32.67	33.25	32.98	32.98
DLM45_01015	CysE	S	27.41	27.00	27.14	27.42	28.56	28.21	28.58	29.13	28.80
DLM45_01255	SoxY	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_01260	SoxZ	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_01870	SoxYZ	S	27.85	27.90	26.77	25.46	26.67	26.38	28.53	29.26	28.61
DLM45_01925	CysH	S	27.87	27.92	27.45	25.07	25.74	25.09	22.73	24.02	25.07
DLM45_01930	CysD	S	30.52	30.62	30.31	29.26	28.83	28.80	23.83	23.36	24.83
DLM45_01935	CysNC	S	30.48	31.47	30.63	28.74	25.53	25.00	24.52	23.17	24.78
DLM45_02930	DmoA	MSC	23.51	25.24	24.56	37.79	37.26	37.15	26.50	25.13	24.36
DLM45_02940	SfnF	MSC	23.64	24.14	23.30	29.72	30.21	30.21	24.15	24.90	24.51
DLM45_02945	SfnG	MSC	29.35	29.55	28.54	32.28	32.84	32.95	29.21	28.89	28.85
DLM45_02965	SoxC	S	27.06	27.31	27.29	31.81	31.22	31.38	26.41	26.94	26.48
DLM45_02970	SoxC	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_03560	MxaF	MeOH	37.68	37.52	37.57	37.07	37.25	36.93	37.78	37.99	37.41
DLM45_03565	MxaJ	MeOH	31.01	30.34	29.44	27.20	29.78	28.58	30.19	31.02	30.75
DLM45_03570	MxaG	MeOH	29.16	28.99	29.13	24.91	26.13	27.51	28.89	29.51	29.02
DLM45_03575	MxaI	MeOH	27.41	25.71	28.17	24.97	23.96	26.51	28.25	28.44	27.23
DLM45_03680	SoxYZ	S	26.00	26.00	25.70	25.85	25.69	26.17	25.17	25.11	25.74

Supplementary Table S5.5-2ad: *Comparative proteomics of* H. methylovorum, *proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			Comparative Analysis								
Protein ID	Protein Name	Pathway	$DMSO_2$ versus MeOH			DMS versus MeOH			$DMSO_2$ versus DMS		
			p-value	Welch's q-value	Fold Change	p-value	Welch's q-value	Fold Change	p-value	Welch's q-value	Fold Change
DLM45_00110	XoxF	MeOH	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_00140	FchB	HCHO	0.125	0.099	0.88	0.092	0.090	1.16	0.006	0.001	0.76
DLM45_00145	FhcA	HCHO	0.529	0.599	1.05	0.671	0.785	1.05	0.999	1.000	1.00
DLM45_00150	FhcD	HCHO	0.190	0.150	0.79	0.571	0.678	1.07	0.084	0.042	0.74
DLM45_00155	FhcC	HCHO	0.420	0.451	0.76	0.751	0.844	1.08	0.334	0.277	0.71
DLM45_00235	MetZ	S	0.161	0.131	1.15	0.017	0.021	1.44	0.044	0.019	0.80
DLM45_01015	CysE	S	0.070	0.080	1.84	0.001	0.005	3.14	0.107	0.068	0.58
DLM45_01255	SoxY	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_01260	SoxZ	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_01870	SoxYZ	S	0.061	0.040	0.39	0.042	0.048	2.44	0.004	0.001	0.16
DLM45_01925	CysH	S	0.001	0.001	0.18	0.005	0.027	0.07	0.127	1.000	1.00
DLM45_01930	CysD	S	0.001	0.003	0.35	0.000	0.005	0.01	0.000	0.000	31.02
DLM45_01935	CysNC	S	0.021	0.036	0.05	0.000	0.003	0.01	0.149	0.103	4.81
DLM45_02930	DmoA	MSC	0.000	0.004	8014.77	0.328	0.331	1.86	0.000	0.001	4311.20
DLM45_02940	SfnF	MSC	0.000	0.000	82.05	0.065	1.000	1.00	0.000	0.000	46.28
DLM45_02945	SfnG	MSC	0.001	0.002	11.67	0.653	0.778	0.90	0.000	0.000	13.03
DLM45_02965	SoxC	S	0.000	0.005	19.04	0.030	0.045	0.66	0.000	0.000	29.05
DLM45_02970	SoxC	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_03560	MxaF	MeOH	0.008	0.015	0.70	0.482	0.583	1.10	0.029	0.015	0.64
DLM45_03565	MxaJ	MeOH	0.117	0.098	0.30	0.490	0.569	1.31	0.053	0.038	0.23
DLM45_03570	MxaG	MeOH	0.018	0.039	0.13	0.815	0.892	1.03	0.019	0.019	0.13
DLM45_03575	MxaI	MeOH	0.134	0.099	0.26	0.344	0.375	1.84	0.027	0.016	0.14
DLM45_03680	SoxYZ	S	0.989	0.994	1.00	0.067	0.075	0.68	0.083	0.038	1.48

Supplementary Table S5.5-2ba: *Comparative proteomics of* H. methylovorum, *proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			Unique Peptides									MS/MS count								
Protein ID	Protein Name	Pathway	MeOH			$DMSO_2$			DMS			MeOH			$DMSO_2$			DMS		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
DLM45_05095	MsmA	MSC	0	0	0	18	18	18	19	20	19	0	0	0	13	12	13	23	21	25
DLM45_05100	MsmB	MSC	0	0	0	3	3	3	3	4	4	0	0	0	5	7	7	4	3	5
DLM45_05105	MsmC	MSC	0	0	0	1	2	2	3	3	2	0	0	0	0	0	0	2	1	2
DLM45_05110	MsmD	MSC	1	1	0	2	2	2	1	1	2	0	1	0	4	1	1	1	1	2
DLM45_07050	Fhs	HCHO	40	39	44	38	35	38	43	41	43	55	68	67	38	35	44	104	99	105
DLM45_07540	FolD	HCHO	19	20	19	13	15	15	21	22	22	34	34	38	11	12	13	43	50	44
DLM45_08590	Pdo	S	6	7	4	2	2	2	20	19	19	0	2	0	0	0	0	40	42	42
DLM45_08925	CysK	S	31	34	29	25	25	22	37	39	37	34	35	32	12	14	12	66	69	63
DLM45_09210	Acs	S	29	29	29	23	22	21	25	25	25	155	178	143	36	30	35	52	43	55
DLM45_09555	MetZ	S	20	21	21	18	19	19	22	18	20	24	25	27	14	12	12	24	16	22
DLM45_11480	catalase	MSC	14	15	12	3	6	2	14	15	15	24	25	10	0	4	2	9	11	10
DLM45_11950	catalase	MSC	32	32	37	36	36	37	55	55	54	54	57	46	49	53	51	178	181	182
DLM45_12160	SoxX	S	0	0	0	0	0	0	4	4	4	0	0	0	0	0	0	5	6	7
DLM45_12165	SoxA	S	1	0	1	7	7	6	17	14	16	0	0	0	2	1	0	17	18	22
DLM45_12175	SoxY	S	0	0	1	0	0	0	4	4	4	0	0	0	0	0	0	16	13	13
DLM45_12180	SoxZ	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_12185	SoxB	S	10	12	11	2	7	4	53	57	58	0	5	3	0	0	0	163	155	167
DLM45_12190	SoxC	S	2	5	2	0	0	0	31	31	31	0	0	0	0	0	0	80	74	80
DLM45_12195	SoxD	S	0	0	0	0	0	0	3	3	3	0	0	0	0	0	0	3	4	4
DLM45_12215	Rhd-SqrB	S	5	6	5	3	2	2	34	33	35	3	3	3	2	1	2	48	46	55
DLM45_12340	MtoX	MSC	37	34	36	22	24	22	62	62	61	39	46	41	9	13	11	503	544	572
DLM45_13300	FolD	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_13340	Fhs	HCHO	4	5	5	4	4	5	2	2	2	9	7	7	10	7	9	0	0	0

Supplementary Table S5.5-2bb: *Comparative proteomics of* H. methylovorum, *proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			Intensity								
Protein ID	Protein Name	Pathway	MeOH			$DMSO_2$			DMS		
			1	2	3	1	2	3	1	2	3
DLM45_05095	MsmA	MSC	0.00	0.00	0.00	2.95E+08	2.38E+08	2.10E+08	5.26E+08	6.23E+08	8.70E+08
DLM45_05100	MsmB	MSC	0.00	0.00	0.00	1.09E+08	9.46E+07	1.10E+08	2.27E+08	2.70E+08	3.39E+08
DLM45_05105	MsmC	MSC	0.00	0.00	0.00	7.07E+06	9.94E+06	1.11E+07	3.52E+07	4.85E+07	4.11E+07
DLM45_05110	MsmD	MSC	2.35E+06	3.64E+08	0.00	3.42E+07	2.78E+07	2.22E+07	2.07E+07	2.70E+07	4.63E+07
DLM45_07050	Fhs	HCHO	1.29E+10	1.40E+10	1.14E+10	4.23E+09	4.17E+09	4.65E+09	1.69E+10	1.79E+10	2.19E+10
DLM45_07540	FolD	HCHO	9.10E+09	7.14E+09	4.50E+09	3.03E+08	5.97E+08	5.83E+08	4.56E+09	5.61E+09	6.30E+09
DLM45_08590	Pdo	S	1.30E+08	1.33E+08	7.30E+07	1.11E+07	1.55E+07	1.37E+07	5.76E+09	7.95E+09	8.59E+09
DLM45_08925	CysK	S	5.16E+09	5.58E+09	4.72E+09	6.19E+08	5.20E+08	4.41E+08	9.34E+09	1.05E+10	1.13E+10
DLM45_09210	Acs	S	8.50E+10	8.78E+10	6.62E+10	3.35E+09	2.88E+09	3.21E+09	9.03E+09	8.54E+09	1.07E+10
DLM45_09555	MetZ	S	5.61E+09	5.31E+09	3.93E+09	4.59E+08	3.10E+08	4.53E+08	1.29E+09	1.06E+09	1.10E+09
DLM45_11480	catalase	MSC	2.42E+09	2.03E+09	6.33E+08	2.49E+07	9.77E+07	2.20E+07	3.94E+08	7.17E+08	6.79E+08
DLM45_11950	catalase	MSC	1.14E+10	1.15E+10	1.02E+10	5.95E+09	5.75E+09	7.14E+09	1.02E+11	9.05E+10	1.00E+11
DLM45_12160	SoxX	S	0.00	0.00	0.00	0.00	0.00	0.00	1.05E+08	1.29E+08	6.63E+07
DLM45_12165	SoxA	S	4.96E+06	0.00	7.01E+06	3.05E+07	4.58E+07	2.93E+07	9.54E+08	9.97E+08	1.11E+09
DLM45_12175	SoxY	S	0.00	0.00	6.60E+06	0.00	0.00	0.00	4.92E+09	4.44E+09	5.79E+09
DLM45_12180	SoxZ	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_12185	SoxB	S	2.39E+08	3.15E+08	2.43E+08	2.15E+08	2.07E+08	1.70E+08	6.60E+10	6.08E+10	7.15E+10
DLM45_12190	SoxC	S	8.27E+06	2.45E+07	3.78E+06	0.00	0.00	0.00	1.44E+10	1.34E+10	1.53E+10
DLM45_12195	SoxD	S	0.00	0.00	0.00	0.00	0.00	0.00	2.63E+08	9.74E+07	4.09E+08
DLM45_12215	Rhd-SqrB	S	1.21E+08	1.54E+08	1.67E+08	5.69E+07	7.17E+07	6.51E+07	2.31E+09	1.78E+09	2.84E+09
DLM45_12340	MtoX	MSC	7.19E+09	6.59E+09	6.64E+09	4.46E+08	8.29E+08	6.64E+08	4.19E+11	4.00E+11	4.32E+11
DLM45_13300	FolD	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_13340	Fhs	HCHO	2.96E+09	3.51E+09	2.88E+09	2.48E+09	2.28E+09	2.12E+09	8.84E+06	1.24E+07	2.12E+07

Supplementary Table S5.5-2bc: *Comparative proteomics of* H. methylovorum, *proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			LFQ Intensity								
Protein ID	Protein Name	Pathway	MeOH			$DMSO_2$			DMS		
			1	2	3	1	2	3	1	2	3
DLM45_05095	MsmA	MSC	25.15	22.90	22.85	30.93	30.44	30.17	28.71	28.79	28.81
DLM45_05100	MsmB	MSC	25.06	23.61	23.66	29.01	29.44	29.30	27.78	27.84	28.13
DLM45_05105	MsmC	MSC	24.07	24.97	24.12	24.95	26.27	25.70	24.54	25.50	24.89
DLM45_05110	MsmD	MSC	23.10	25.49	23.87	27.79	27.39	26.98	24.10	24.64	25.10
DLM45_07050	Fhs	HCHO	33.87	34.14	34.06	33.96	33.83	33.95	34.38	34.45	34.38
DLM45_07540	FolD	HCHO	33.03	32.87	32.48	31.39	31.85	31.79	31.85	32.16	32.16
DLM45_08590	Pdo	S	25.56	25.69	24.98	24.94	25.35	25.50	32.48	32.71	32.51
DLM45_08925	CysK	S	32.25	32.53	32.56	32.00	31.80	31.61	32.96	33.04	32.91
DLM45_09210	Acs	S	36.09	36.46	36.38	34.26	33.99	34.05	33.18	32.97	33.09
DLM45_09555	MetZ	S	32.34	32.47	32.37	31.39	31.24	31.33	30.10	29.92	29.82
DLM45_11480	catalase	MSC	31.10	31.16	29.70	27.68	29.09	27.48	28.64	29.30	28.63
DLM45_11950	catalase	MSC	33.37	33.59	33.81	35.04	34.87	34.94	36.45	36.20	36.06
DLM45_12160	SoxX	S	24.19	23.64	24.13	26.73	25.41	25.30	26.50	26.68	26.14
DLM45_12165	SoxA	S	24.20	26.56	25.16	27.97	28.40	27.75	29.48	29.60	29.55
DLM45_12175	SoxY	S	24.88	25.74	24.14	25.20	25.65	24.65	32.03	31.86	31.92
DLM45_12180	SoxZ	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_12185	SoxB	S	25.73	27.34	25.95	25.53	26.73	24.59	36.03	35.80	34.65
DLM45_12190	SoxC	S	20.62	21.80	20.21	25.18	23.31	25.89	33.55	33.39	33.44
DLM45_12195	SoxD	S	23.66	23.38	24.89	26.46	24.21	26.84	27.70	27.60	27.87
DLM45_12215	Rhd-SqrB	S	26.80	27.43	27.70	28.75	28.92	28.68	30.88	30.44	31.06
DLM45_12340	MtoX	MSC	32.52	32.85	33.10	31.50	32.20	31.66	38.34	38.31	38.11
DLM45_13300	FolD	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_13340	Fhs	HCHO	31.74	31.89	32.21	32.63	34.15	33.82	23.30	23.67	23.42

Supplementary Table S5.5-2bd: *Comparative proteomics of* H. methylovorum, *proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			Comparative Analysis								
Protein ID	Protein Name	Pathway	$DMSO_2$ versus MeOH			DMS versus MeOH			$DMSO_2$ versus DMS		
			p-value	Welch's q-value	Fold Change	p-value	Welch's q-value	Fold Change	p-value	Welch's q-value	Fold Change
DLM45_05095	MsmA	MSC	0.001	0.009	118.05	0.003	0.024	35.19	0.001	0.004	3.35
DLM45_05100	MsmB	MSC	0.000	0.009	35.36	0.001	0.015	13.98	0.001	0.001	2.53
DLM45_05105	MsmC	MSC	0.059	0.042	2.39	0.215	0.197	1.51	0.234	0.154	1.58
DLM45_05110	MsmD	MSC	0.012	0.024	9.43	0.576	0.703	1.38	0.002	0.001	6.84
DLM45_07050	Fhs	HCHO	0.295	0.295	0.93	0.010	0.030	1.30	0.000	0.001	0.71
DLM45_07540	FolD	HCHO	0.007	0.009	0.46	0.019	0.027	0.60	0.104	0.051	0.77
DLM45_08590	Pdo	S	0.610	0.689	0.90	0.000	0.002	142.66	0.000	0.000	0.01
DLM45_08925	CysK	S	0.013	0.012	0.64	0.008	0.025	1.44	0.001	0.001	0.45
DLM45_09210	Acs	S	0.000	0.007	0.22	0.000	0.004	0.11	0.001	0.000	2.03
DLM45_09555	MetZ	S	0.000	0.000	0.47	0.000	0.003	0.18	0.000	0.000	2.60
DLM45_11480	catalase	MSC	0.021	0.018	0.17	0.027	0.042	0.29	0.233	0.177	0.58
DLM45_11950	catalase	MSC	0.001	0.006	2.57	0.000	0.003	6.27	0.000	0.001	0.41
DLM45_12160	SoxX	S	0.021	1.000	1.00	0.000	0.001	5.46	0.267	0.226	0.65
DLM45_12165	SoxA	S	0.018	0.033	6.65	0.003	0.026	18.86	0.002	0.004	0.35
DLM45_12175	SoxY	S	0.671	1.000	1.00	0.000	0.007	129.81	0.000	0.001	0.01
DLM45_12180	SoxZ	S	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_12185	SoxB	S	0.417	0.428	0.61	0.000	0.003	569.94	0.000	0.000	0.00
DLM45_12190	SoxC	S	0.012	0.015	15.09	0.000	0.005	6150.81	0.000	0.002	0.00
DLM45_12195	SoxD	S	0.120	1.000	1.00	0.001	0.016	13.40	0.084	0.078	0.27
DLM45_12215	Rhd-SqrB	S	0.006	0.021	2.77	0.000	0.002	11.18	0.001	0.000	0.25
DLM45_12340	MtoX	MSC	0.019	0.018	0.49	0.000	0.003	43.26	0.000	0.000	0.01
DLM45_13300	FolD	HCHO	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_13340	Fhs	HCHO	0.030	0.042	3.01	0.000	0.000	0.00	0.000	0.001	1077.05

Supplementary Table S5.5-2ca: *Comparative proteomics of* H. methylovorum, *proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			Unique Peptides									MS/MS count								
Protein ID	Protein Name	Pathway	MeOH			$DMSO_2$			DMS			MeOH			$DMSO_2$			DMS		
			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
DLM45_14260	XoxF	MeOH	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_14665	SoxX	S	0	0	0	0	0	0	3	3	3	0	0	0	0	0	0	4	5	4
DLM45_14670	SoxY	S	0	0	0	1	0	1	4	4	4	0	0	0	0	0	0	10	9	11
DLM45_14675	SoxZ	S	2	0	3	2	3	4	6	6	6	0	0	0	0	1	1	32	39	35
DLM45_14680	SoxA	S	2	1	1	0	2	0	9	10	10	1	0	0	0	0	0	22	18	17
DLM45_14685	SoxB	S	14	18	16	12	11	15	45	45	45	6	4	2	0	0	2	132	127	139
DLM45_14690	SoxC	S	1	3	2	0	1	1	21	21	22	0	0	0	0	0	0	40	33	38
DLM45_14695	SoxC	S	0	0	0	0	0	0	4	5	5	0	0	0	0	0	0	4	5	3
DLM45_16005	Fae	HCHO	11	10	11	9	11	11	12	13	13	65	39	41	28	40	49	69	95	118
DLM45_16020	Mch	HCHO	18	18	18	13	13	14	18	19	18	16	16	15	13	11	11	36	34	35
DLM45_16030	MtdB	HCHO	17	17	16	16	16	17	18	18	19	58	52	42	27	34	35	58	63	66
DLM45_16080	CysJ	S	24	23	21	2	2	2	0	0	0	28	30	22	3	3	2	0	0	0

Supplementary Table S5.5-2cb: *Comparative proteomics of* H. methylovorum, *proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			Intensity								
Protein ID	Protein Name	Pathway	MeOH			$DMSO_2$			DMS		
			1	2	3	1	2	3	1	2	3
DLM45_14260	XoxF	MeOH	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_14665	SoxX	S	0.00	0.00	0.00	0.00	0.00	0.00	6.69E+07	3.27E+07	8.70E+07
DLM45_14670	SoxY	S	0.00	0.00	0.00	3.62E+05	0.00	3.56E+06	2.78E+09	2.63E+09	3.14E+09
DLM45_14675	SoxZ	S	1.32E+07	0.00	1.19E+07	3.59E+06	1.30E+07	2.61E+07	4.15E+09	5.86E+09	6.60E+09
DLM45_14680	SoxA	S	2.08E+07	1.54E+07	7.07E+06	0.00	6.61E+06	0.00	6.59E+08	9.11E+08	1.01E+09
DLM45_14685	SoxB	S	2.96E+08	3.10E+08	3.51E+08	9.18E+07	7.62E+07	1.08E+08	2.91E+10	3.04E+10	3.55E+10
DLM45_14690	SoxC	S	5.34E+07	1.40E+08	3.90E+07	0.00	1.02E+07	1.19E+07	2.28E+09	2.43E+09	3.07E+09
DLM45_14695	SoxC	S	0.00	0.00	0.00	0.00	0.00	0.00	1.26E+08	1.39E+08	1.71E+08
DLM45_16005	Fae	HCHO	7.04E+10	3.93E+10	4.80E+10	1.13E+10	3.09E+10	2.89E+10	3.41E+10	4.54E+10	5.05E+10
DLM45_16020	Mch	HCHO	6.23E+09	3.61E+09	2.41E+09	5.69E+08	7.05E+08	7.57E+08	4.38E+09	4.93E+09	6.08E+09
DLM45_16030	MtdB	HCHO	9.25E+09	7.35E+09	4.83E+09	1.89E+09	2.59E+09	2.65E+09	6.06E+09	7.44E+09	9.60E+09
DLM45_16080	CysJ	S	2.18E+09	2.24E+09	1.79E+09	1.00E+08	1.46E+08	1.30E+08	0.00	0.00	0.00

Supplementary Table S5.5-2cc: *Comparative proteomics of* H. methylovorum, *proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			LFQ Intensity								
Protein ID	Protein Name	Pathway	MeOH			$DMSO_2$			DMS		
			1	2	3	1	2	3	1	2	3
DLM45_14260	XoxF	MeOH	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_14665	SoxX	S	24.78	24.11	24.81	24.61	24.66	23.98	25.69	25.73	25.55
DLM45_14670	SoxY	S	24.42	25.77	23.39	25.81	25.29	26.16	31.27	31.02	31.07
DLM45_14675	SoxZ	S	21.85	24.70	22.33	22.56	24.24	24.37	31.61	32.06	31.83
DLM45_14680	SoxA	S	24.45	24.32	23.76	25.86	25.25	24.73	29.61	29.51	28.46
DLM45_14685	SoxB	S	28.35	28.74	28.17	28.01	27.54	28.57	34.44	34.58	34.65
DLM45_14690	SoxC	S	25.76	26.57	25.44	25.41	23.89	24.79	30.98	31.00	31.00
DLM45_14695	SoxC	S	23.97	24.92	25.05	25.56	25.20	25.47	26.93	26.84	27.06
DLM45_16005	Fae	HCHO	35.67	35.04	35.92	35.80	37.21	37.35	35.20	35.40	35.34
DLM45_16020	Mch	HCHO	32.27	31.95	31.60	31.43	31.75	31.65	32.13	32.23	32.27
DLM45_16030	MtdB	HCHO	32.80	32.72	32.23	32.46	32.80	32.67	32.56	32.84	32.88
DLM45_16080	CysJ	S	30.90	31.22	31.13	29.55	29.92	29.68	23.56	24.44	22.34

Supplementary Table S5.5-2cd: *Comparative proteomics of* H. methylovorum, *proteins of interest intensity data. Unidentified proteins are marked as N/A.*

Enzyme			Comparative Analysis								
Protein ID	Protein Name	Pathway	$DMSO_2$ versus MeOH			DMS versus MeOH			$DMSO_2$ versus DMS		
			p-value	Welch's q-value	Fold Change	p-value	Welch's q-value	Fold Change	p-value	Welch's q-value	Fold Change
DLM45_14260	XoxF	MeOH	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLM45_14665	SoxX	S	0.662	1.000	1.00	0.010	0.032	2.13	0.005	0.009	0.42
DLM45_14670	SoxY	S	0.169	0.169	2.34	0.001	0.014	96.63	0.000	0.000	0.02
DLM45_14675	SoxZ	S	0.509	0.564	1.70	0.001	0.012	468.87	0.000	0.000	0.00
DLM45_14680	SoxA	S	0.046	0.038	2.15	0.000	0.004	32.38	0.001	0.001	0.07
DLM45_14685	SoxB	S	0.326	0.324	0.77	0.000	0.003	70.32	0.000	0.001	0.01
DLM45_14690	SoxC	S	0.092	0.069	0.43	0.000	0.007	33.65	0.000	0.001	0.01
DLM45_14695	SoxC	S	0.099	1.000	1.00	0.003	0.021	4.91	0.000	0.000	0.35
DLM45_16005	Fae	HCHO	0.090	0.082	2.37	0.437	0.530	0.85	0.041	0.040	2.78
DLM45_16020	Mch	HCHO	0.201	0.184	0.80	0.241	0.286	1.21	0.004	0.003	0.66
DLM45_16030	MtdB	HCHO	0.787	0.856	1.04	0.437	0.497	1.13	0.454	0.446	0.92
DLM45_16080	CysJ	S	0.001	0.003	0.39	0.000	0.008	0.01	0.001	0.002	77.12

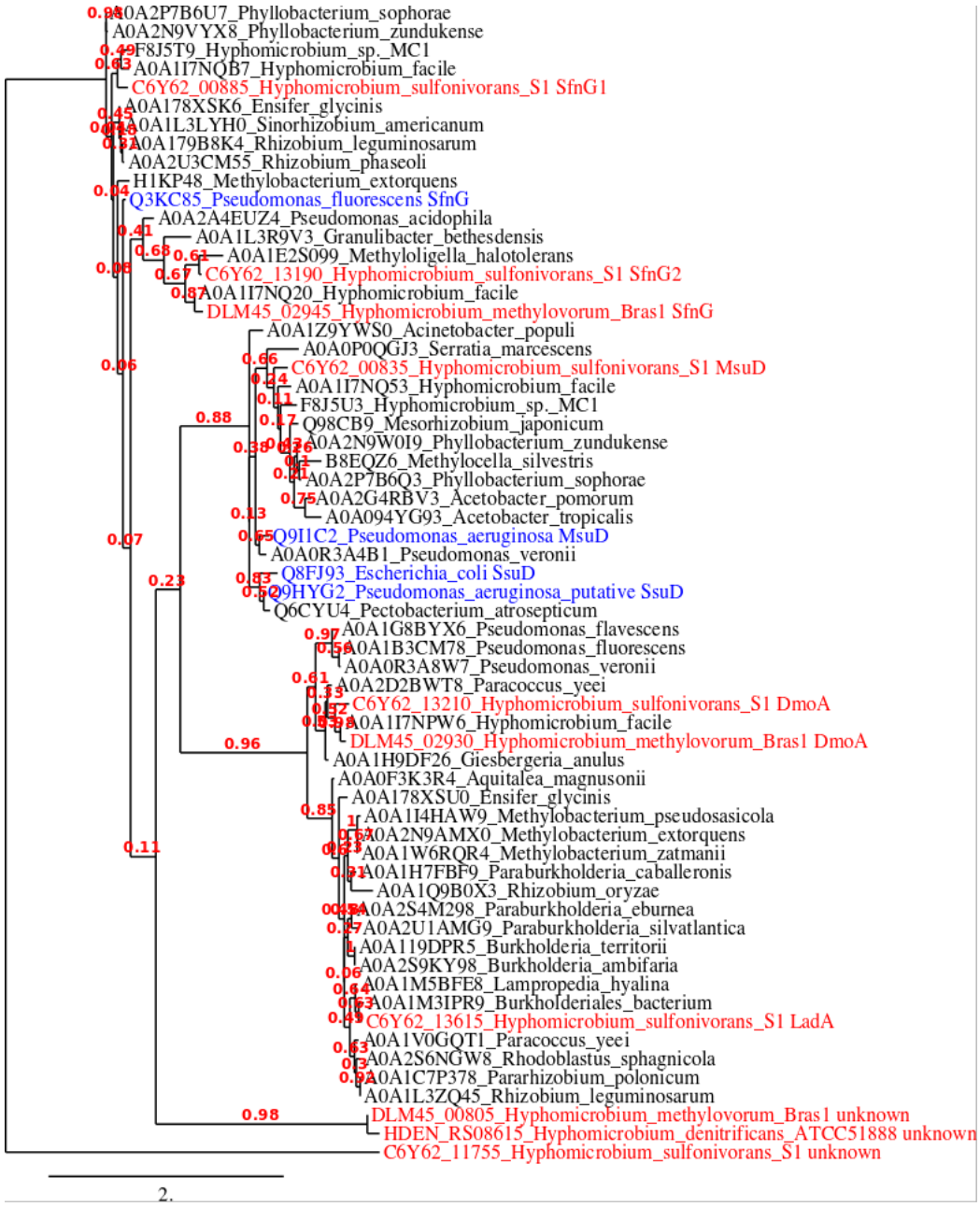


Figure S5.6-1: *Phylogenetic tree of $FMNH_2$-dependent monooxygenase large subunits in* Hyphomicrobium *species, against $FMNH_2$-dependent monooxygenase-like sequences from other bacterial species.* Experimentally characterised reference sequences are labelled BLUE, while sequences derived from *H. denitrificans* ATCC51888, *H. methylovorum* Bras1 and *H. sulfonivorans* S1 are labelled RED.